An effective approach for mapping permafrost in a large area using subregion maps and satellite data

Jianan Hu1 | Shuping Zhao 1 | Zhuotong Nan1,2 | Xiaobo Wu3 | Xuehui Sun1 | Guodong Cheng4

Abstract
Permafrost distribution maps are of importance for environmental assessment, climate system modeling, and practical engineering applications. The scarcity of forcing data and parameters often limits the uses of permafrost models over large areas. However, detailed data are often available in a few subregions through field investigations. In this study, we propose a novel approach for mapping permafrost distribution in a large and data-scarce area using an empirical model with subregion permafrost maps and satellite data as inputs. The surface frost number model (FROSTNUM) was re-inferred to include an extra soil parameter to represent the thermal and moisture conditions in soils. The optimal soil parameters were determined from the subregion maps of permafrost distribution through spatial clustering, parameter optimization, and the decision tree method. FROSTNUM was fed with satellite-derived ground surface freezing and thawing indices to map the permafrost distribution over the study area. The proposed approach was evaluated in the Gaize area on the Qinghai–Tibet Plateau, where intensive field studies have been done. The simulated permafrost distribution is consistent with a map of permafrost distribution made from borehole observations and field surveys in Gaize. Due to excellent accuracy, the approach is effective and can be used in large areas with limited data.

1 | INTRODUCTION

The permafrost distribution map is of great utility for cold region engineering applications or as a basis for assessing the ecological environment and developing proactive adoptions to climate change. Improving the accuracy of permafrost mapping has significant scientific and practical significance.1–3 Traditional permafrost mapping approaches involve costly field investigations that are generally time-consuming, labor-intensive, and potentially unreliable if for a large area with complex conditions.4 In recent decades, permafrost models at a variety of complexity have been extensively used for mapping permafrost distribution.4–7

Permafrost models can be categorized into empirical/semi-physical models and physical models. Given that the distribution of permafrost is related to many external factors such as locations, elevation, and air temperature, many empirical models have been
developed through establishing relationships between the distribution and those factors, such as the elevation model,5 mean annual ground temperature (MAGT) approach,9 the temperature at the top of permafrost (TTOP) model,10 and the frost number model.13 These models are still in extensive use on data-scarce regions such as the Qinghai-Tibet Plateau (QTP) due to model simplicity and minimal data requirements.12–16 However, empirical models assume steady-state conditions and oversimplify the account of soil heterogeneity. They tend to be incapable of reflecting inhomogeneity in local factors that modulate the hydrological and thermal processes within frozen soils. As a result, empirical models that do not consider the impacts of local factors such as topography, soil texture, and soil moisture condition may lead to large discrepancies in mapping permafrost distribution.16

Zhao et al. have shown that the performance of empirical models can be improved by explicitly considering the effects of local factors.5 On the other hand, physical models are theoretically superior to empirical models with regard to their ability to quantify hydrothermal processes occurring in frozen ground. The impacts of local factors on permafrost formation and distribution can be represented by spatially distributed parameters in the model. Physical models have been applied to some site-scale studies on the QTP (such as applications of the Community Land Model,17 Common Land Model,18 Simultaneous Heat and Water Model,19 Coup Model,20 and the Noah land surface model21) and small areas with rich investigation data.22 They have been rarely used in a large area with limited data due to their rigid requirements of data and model parameters. In a specific application, the choice of a proper permafrost model depends greatly on the availability of the data and model parameters in the study area. In a data-scarce area, it is practical to consider the empirical/semi-physical model as an alternative for a permafrost mapping effort.

The inputs for a permafrost model include forcing data providing upper and lower boundaries to the model, and model parameters such as soil and vegetation conditions. In a large area such as the QTP, observation sites are minimal. There are about 80 meteorological stations and dozens of permafrost monitoring sites in a total area of 2.6 million square kilometers. Most of the sites are mainly along the Qinghai-Tibet Highway and few in the western part of the QTP. Therefore, it is impossible to provide adequate observation data and model parameters for mapping the entire region. Satellite-derived data and re-analysis data can compensate for the shortage of forcing data. As remote sensing provides globally or regionally covered, uninterrupted observations at high spatial, spectral, and temporal resolutions, the derived and re-analysis data are presently widely used as the upper boundaries of permafrost models.23–26 Model parameters such as soil properties, however, are more dependent on field measurements and are difficult to acquire with indirect approaches. This will introduce considerable uncertainty to the results if sensitive parameters that exhibit strong horizontal and vertical heterogeneity are not properly assigned. In some previous studies, due to the lack of reliable soil parameters, the effects of soil heterogeneity on permafrost have been neglected, leading to compromised reliability in the results.

Recently, several permafrost investigation projects have been initiated in China, the United States, Canada, Russia, and pan-Arctic countries.27–31 Through these efforts, many data have been collected from field survey and laboratory analysis. Since 2009, a China Minister of Science and Technology sponsored 5-year project was carried out to investigate permafrost and its environment over the QTP. Five representative areas (West Kunlun, Gaize, Wenquan, Aerjin, and National Highway 308) were chosen to perform intensive surveys.27 The second scientific expeditions on the QTP have already collected and produced more field data for the region with extensive underlying permafrost.28 Nevertheless, for such a large area like the QTP, investigations remain concentrated in relatively small regions and long-lasting efforts are needed to improve the data conditions.

Because the investigation data represent the local conditions with good accuracy, the local permafrost distribution maps based on survey data are generally of good quality. In this paper, we fully mine the underlying information in the subregion maps of permafrost distribution that are of reliable quality, in order to create a high-accuracy permafrost map for a large study area encompassing those subregions through a semi-physical model, which takes into account the effects of local factors, along with remotely sensed surface temperature data. The approach is evaluated in a transition zone from permafrost to seasonally frozen ground, the Gaize area on the QTP, to be applicable and effective.

2 METHODOLOGY

2.1 The extended surface frost number model

The surface frost number model (FROSTNUM)31 was developed to determine the occurrence of permafrost quantified by a frost number, which is calculated from the freezing and thawing indices of ground surface temperature (GST). We extended FROSTNUM to include an extra parameter related to soil hydrothermal properties.32

The depth of thaw, \( \xi_t \), in the thawing season, and the depth of freeze, \( \xi_f \), in the freezing season, at any location of frozen ground can be expressed as:

\[
\xi_t = \alpha \xi_f
\]  

(1)

where \( \alpha \) is the ratio of the depth of thaw over the depth of freeze; \( \alpha > 1 \) represents a seasonally frozen ground case and \( \alpha \leq 1 \) a permafrost case.

Equation 1 can be reorganized into:

\[
\xi_t + \xi_f = (\alpha + 1)\xi_f
\]  

(2)

\[
\frac{1}{\alpha + 1} = \frac{\xi_f}{\xi_t + \xi_f}
\]  

(3)

Assuming no convective heat transfer has occurred and there is a uniform soil texture, the freeze and thaw depths can be analytically
derived by Stefan’s formula given an initial ground surface temperature of 0°C:

$$\xi_f = \sqrt{\frac{-2q_{DF}}{Q_t}}$$  \hspace{1cm} (4)

$$\xi_t = \sqrt{\frac{-2q_{DT}}{Q_t}}$$  \hspace{1cm} (5)

where DDF and DDT are the freezing and thawing indices (°C.day), respectively, on the ground surface; \(\lambda\) is soil thermal conductivity (W m\(^{-1}\) K\(^{-1}\)); and Q is heat released or absorbed by a unit volume of soil (kJ m\(^{-3}\)) during freezing or thawing periods, \(Q = L \cdot \gamma \cdot (W - W_u)\). \(L\) is the latent heat of fusion for ice and liquid (3.3 \times 10^5 \text{ kJ m}^{-3}). \(W\) and \(W_u\) are total and unfrozen soil water contents on a gravimetric basis (%), respectively, \(\gamma\) is the dry density of soil (kg m\(^{-3}\)). Subscripts \(f\) and \(t\) denote frozen and unfrozen soils, respectively.

Let \(F\) defined as the frost number, denote \(1/(\alpha + 1)\); we obtain \(F\) by substituting Equations 4 and 5 into Equation 3:

$$F = \frac{\sqrt{DDF}}{\sqrt{DDF} + E \cdot \sqrt{DDT}}$$  \hspace{1cm} (6)

$$E = E_t \cdot Q_t\sqrt{\frac{\xi_t}{\xi_f}}$$  \hspace{1cm} (7)

where \(E\) is a dimensionless parameter representing the changes in soil properties from the unfrozen state to the frozen state and is determined by soil thermal properties and moisture conditions on both states. For the seasonally frozen soil, the thaw depth would be larger than freeze depth, i.e., \(\alpha > 1\) and \(F < 0.5\). For permafrost, \(F > 0.5\). It is physically meaningful to take \(F = 0.5\) as a threshold for identifying permafrost and seasonally frozen ground in Equation 6.

In an idealized case ignoring the occurrence of soil property changes during phase transition and the effect of convection on the formation of frozen soils, \(E = 1\), and the original Nelson’s frost number model is then obtained:

$$F = \frac{\sqrt{DDF}}{\sqrt{DDF} + \sqrt{DDT}}$$  \hspace{1cm} (8)

However, under realistic conditions, frozen soils generally have higher thermal conductivity than unfrozen soils. Therefore, a strong thermal offset can be produced. It preserves permafrost even under conditions where ground freezing indices are lower than the thawing indices. Similarly, the existence of convection of liquid water in soil strata may result in no permafrost even though the freezing indices are greater than the thawing indices. In this sense, if Equation 8 is used, \(F = 0.5\) as a threshold for identifying frozen soil types is no longer physically correct and the threshold must be modified to meet the real situation. By introducing an extra soil-related parameter into the model, Equation 6 enables more flexibility and allows for more space for parametric optimization even if the parameter cannot be measured.

3 | OVERALL SCHEME FOR THE FROSTNUM/COP APPROACH

The FROSTNUM/COP requires high-quality subregion maps of permafrost distribution within the study area to inversely compute the soil parameter \(E\) using spatial clustering and parametric optimization techniques. The soil parameters are then determined for the entire study area using a classification method (i.e., the decision tree). As model inputs, ground surface freezing and thawing indices are estimated from remotely sensed land surface temperature (LST) data. The overall scheme of this approach is described in Figure 1. It includes the following steps:

1 Preparing freezing and thawing indices. As permafrost is a product of long-term climatic conditions, we acquired multiyear LST data rather than from a single year. Although we collected data in 2003–2010 in the case study, a longer data series, if available, will be better representative of local temperature conditions. The daytime and nighttime LST data are each decomposed using wavelet transforms. The adaptive network-based fuzzy inference system (ANFIS) is then used to estimate daily GST. After properly interpolating the missing mean daily GSTs, the surface ground freezing and thawing indices on each modeling cell are computed.

2 Clustering and optimizing soil parameters from subregion maps. In the subregions where permafrost distribution maps are available, the soils are spatially clustered according to a variety of local factors. The soil parameter \(E\) in the extended FROSTNUM is then estimated for each soil cluster by the particle swarm optimization (PSO) method. PSO takes the subregion maps as the optimization objective and iterates the model to find optimal parameter values.

3 Determining soil cluster distribution over the entire study area. Using the distribution of soil clusters in the subregions as a training set, a decision tree method (i.e., C50) is adopted to determine the distribution of soil clusters on unknown locations. Variables including topography, soil texture, and soil moisture condition serve as the predictors.

4 Mapping the permafrost distribution for the study area. Fed by freezing and thawing indices along with the optimally determined soil parameters, the extended FROSTNUM model is run on all modeling cells to calculate frost numbers, upon which the type of frozen soil is identified by a threshold method.

3.1 | Preparing freezing and thawing indices

3.1.1 | Estimating daily GST from LST

Meteorological stations record 0 cm soil temperature but they are usually sparse and unevenly distributed particularly in a large area.
Therefore, a regional-scale distribution of GST is generally difficult to obtain by ground observations. Remotely sensed data such as Moderate Resolution Imaging Spectroradiometer (MODIS) land data products offer a possible way to estimate regional GSTs with global coverage at high spatial and temporal resolutions.

GST exhibits high spatial heterogeneity and is nonlinearly related to the LST. A Wavelet-ANFIS approach proposed by Huang et al. is recommended to estimate daily GST from two satellite LST observations. Inputs include daytime and night-time MODIS LSTs, which will be decomposed into one low pass and three high passes each by a wavelet transform, and the time of sunrise at the given day and location to be processed. The wavelet function is set to reverse biorthogonal 3.1. As many as possible in situ daily GST records in the study area are collected for training ANFIS to maximize prediction performance. In the case that the study area has too few GST observations, a larger area encompassing the study area can be used to ensure sufficient GST observations. The desired GSTs for the study area can be extracted later.

### 3.1.2 Aggregating multiyear daily GSTs and calculating freezing and thawing indices

Ideally, for an \( m \) year span, there should be \( m \) GST values in a day of the year at every location. However, MODIS LST data are subject to heavy cloud contamination in some regions, especially on the QTP. The actual situation may be worse because some daily GST values may not be available in the absence of valid MODIS LSTs for that day. To ensure quality, only if there are at least three valid GST values for a given day of year is the daily multiyear mean GST arithmetically aggregated for that day. Otherwise, it is treated as missing at this location.

A cosine interpolation, which can accurately reproduce the annual oscillation of temperature, is used to fill in the missing GSTs if the number of valid days in a location is more than 120. The remainder, which probably accounts for a small portion, is spatially interpolated using Ordinary Kriging, which can satisfactorily provide year-around daily mean GSTs for the entire study area.

Based on the daily mean GSTs across multiple years, the annual freezing indices for each cell are calculated, per definition, as the sum of degrees that a day's GST is below 0°C. Likewise, the thawing indices are counted as the sum of degrees that a day's GST is above 0°C.

### 3.2 Inferring soil parameters from subregion maps

#### 3.2.1 Spatial clustering

The soil parameter (\( E \)) in the extended FROSTNUM is related to soil hydrothermal conditions. Soils that have similar hydrothermal conditions are likely to share the same parametric value. As environmental factors play critical roles in controlling the soil distribution and associated hydrothermal properties, the topographic factors, soil texture, and soil moisture variables are selected as attributes to group the soils across the subregions using a spatial clustering technique. Soils in a cluster share more similarity in hydrothermal characteristics to each other than to those in other clusters. But a soil cluster is not necessarily a specific soil type defined in soil science. The k-prototype approach is used as there are both numerical and categorical...
variables involved. The number of clusters needs to be iteratively adjusted to best represent the real distribution of soil conditions over the subregions. Generally, the cluster number is preset to the number of predominant soil textures in a local soil map. It is then trialed by evaluating the agreement of the simulated permafrost distribution against the subregion maps. An optimal number can be singled out from candidate numbers to give good agreement.

3.2.2 | Parametric optimization

As there is no direct way to measure the soil parameter $E$, a parametric estimation technique can be used to obtain optimal parameter values for soil clusters, provided that frozen soil distribution maps are available in some subregions. Those subregion maps are usually of relatively high quality as they were created with adequate field data and adjusted with expertise. PSO is prescribed because this population-based metaheuristic method can converge to an optimal value quickly. The Kappa coefficient is used as the objective function to measure the consistency between the simulation results and the subregion permafrost distribution maps. The optimal values of $E$ associated with soil clusters can be found after intensive computations and a junction table is subsequently created to look up the type of soil cluster in exchange for the $E$ value.

3.3 | Mapping soil cluster distribution and permafrost distribution over the study area

With an assumption of parametric stationarity, the optimized soil parameter values are transferred to the entire study area. A decision tree approach is used as the transfer function to extract rules associating the environmental factors with the type of soil cluster. More specifically, the distribution of soil clusters across the subregions is provided to the C50 classifier as a training set, and the rules are then applied to predict the soil cluster types at unknown locations. The soil parameter values are determined by looking up the junction table established at an early stage. Finally, the extended FROSTNUM model is performed on each grid cell. The type of frozen soil is identified by a threshold frost number of 0.5. The permafrost cells have frost numbers $\geq 0.5$ and the seasonally frozen soil cells have numbers $<0.5$. The locations where the frost numbers equal zero (i.e., the freezing indices are equal to zero) are assigned a nonfrozen type.

3.4 | Experimental study in the Gaize area, Tibet

We selected Gaize (Figure 2) as an experimental study area for evaluating the FROSTNUM/COP approach. The area is representative of a thermally unstable permafrost area, where local factors strongly influence the distribution of permafrost. A high-quality regional permafrost map is available for Gaize. Three subregions (R1, R2, and R3 in Figure 2) were selected in the areas with numerous boreholes. We extracted the distributions of frozen ground type in the three subregions from the permafrost map of Gaize. The three subregion maps of permafrost distribution are expected to be of higher precision than the other areas due to the presence of adequate field information.

We used the FROSTNUM/COP approach to calibrate the soil parameters based on the subregion maps. Finally, the FROSTNUM/COP approach produced a new map of the entire Gaize region, which was then assessed against the existing permafrost map of Gaize and a recently published map from Zou et al. The Gaize area is bounded within longitudes 84°–86.5°E and latitudes 32°–34°N (Figure 2), located in a transition zone between alpine permafrost and seasonally frozen ground, covering an area of...


\[ 41.2 \times 10^3 \text{ km}^2 \]. Elevations in the Gaize area vary from 4,400 to 6,200 m above sea level (a.s.l.), with an average of 4,700 m a.s.l. Alpine steppe is the primary vegetation type in the region. Annual average temperature is around 0°C. Annual average precipitation is about 150 mm, and 90% of the precipitation is concentrated in the warm season of May to September. Boreholes in the area reveal that the typical soil is composed of coarse sand and gravel. The parent materials in the study area are mainly clastic rocks, proluvium, and lacustrine sediments, and are generally characterized by thin layers of gravel or with an abundance of calcium carbonate and soluble salts. Permafrost tables were detected at a depth below 3.5 m and the mean annual ground temperatures are primarily above \(-1.5^\circ\text{C}\)\(^{41}\).

A substantial amount of fieldwork was conducted in the Gaize area in 2010. It includes 20 test pits, 22 boreholes (8–56 m) as marked in Figure 2 (some boreholes overlap on the figure), 14 traverses of ground-penetrating radar, and 22 traverses of Geophysical Data Processor Receivers.\(^{42}\) The data provide the basis for the regional permafrost map of Gaize. The map was made based on regression analyses of the altitudinal limit of permafrost (i.e., the lowest elevation that permafrost occurs) that were performed for various hillslope aspects, and further improved by incorporating expert knowledge such as regarding the impacts of roads and rivers, lakes and wetlands, and periglacial geomorphology across the area. The permafrost map of Gaize is provided at a resolution of 250 m and was upscaled to 1 km to match the model resolution.

There is only one meteorological station for Gaize County providing observed GSTs, which is insufficient for training the Wavelet-ANFIS model. Therefore, we performed the Wavelet-ANFIS model for the entire QTP, where a total of 71 meteorological stations (Figure 2) are available and recorded GST at least four times a day throughout 2003–2010. They were aggregated to daily mean GST as training data. The MODIS/Aqua LST production MYD11A1 version 6 (daily, 1 km) data from 2003 to 2010 were processed and used for estimating daily mean GSTs. The Aqua satellite crosses the Equator at about 1:30 p.m. and 1:30 a.m. local time. The topographical factors, including elevation, slope, aspect, and relief degree, were derived from 90-m land elevation data from the Shuttle Radar Topography Mission (SRTM/DEM) and were resampled to 1 km. Hillslope aspects were reclassified into three types: north, south, and east–west. The soil texture data were a subset from a 1-km soil properties dataset covering the mainland of China.\(^{43}\) The soil moisture condition factors include topographic wetness index (TWI), which was also derived from SRTM/DEM, and mean annual total precipitation, which was calculated from the daily and 8-km CMORPH bias-corrected data in the same period. In our testing, an 8-year span of precipitation data are sufficient to give an average pattern of precipitation in the area.

When performing parameter optimization, the extended FROSTNUM model was used as the model operator. The range of \(E\) was set to 0.5–1.5. Iteration times and population size were set to 1,000 and 20, respectively. The number of clusters was determined as five after multiple trials. Both Cohen's Kappa coefficient (\(\kappa\))\(^{44}\) and overall accuracy (OA)\(^{45}\) were used to evaluate the performance of this approach. The former is generally thought to be a more robust measure than OA in accounting for chance agreement, although sometimes it also suffers from the penalty associated with chance agreement.\(^{46}\)

## RESULTS AND DISCUSSION

### 4.1 Freezing and thawing indices

The fractions of invalid MODIS LST data and the aggregated GST estimates on the QTP throughout 2003–2010 are shown in Figure 3. The invalid fractions are calculated as the number of invalid pixels divided by total pixels. The invalid fractions of the LST data are calculated for each day of the year as the averages over the 8 years. There were a larger number of invalid LST pixels on the plateau each day because of heavy cloud contamination. The proportions varied from 3% to 67%. On average, there were 28.4% invalid LST pixels during daytime and 7.8% during nighttime in 2003–2010. In the warm season (May to October), the average proportions of invalid LST pixels were 32.94% in the daytime and 10.7% in the nighttime, while in the cold season (November to April), values were 24.2% in the daytime and 4.8% in the nighttime. The night-time and cold seasons had less missing LST observations than the daytime and warm seasons on the QTP. The aggregated daily GST data suffered severe losses as well because we flagged as missing those with fewer than three valid GST estimates on a day of the year. The situation was improved significantly by the temporal cosine interpolation. The average percentage of invalid GST pixels decreased from 31% to 5%, as shown by the red lines in Figure 3. The remaining 5% missing was then completed by the Ordinary Kriging method.

The quality of the daily GST estimates was verified against the observed GSTs as shown in Figure 4. The 8-year (2003–2010) observed daily mean GST data at 60 weather stations were used as training samples to the Wavelet-ANFIS approach, and the records for the remaining 11 stations were used as validation data. Daily GST estimates were in agreement with the observed data, with the linear trend lines close to 1:1 for both the training set and the validation set. The multiple correlation coefficients (\(R^2\)) were up to 0.90 in the training stage and 0.91 in the validation stage, and the root-mean-square errors (RMSEs) were 2.64 and 2.55°C, respectively. These metrics show that estimating GST from LST is applicable and acceptable for predictive purposes. To achieve better accuracy, the GST records from all available sites were input into the approach to train a final model, by which the daily mean GST data we used in this study were estimated.

The thawing and freezing indices in the Gaize area were calculated from the aggregated daily GST estimates. The spatial patterns of the thawing and freezing indices were in close agreement with the elevation distribution. The thawing indices in Gaize ranged from 541 to 3510°C-d (Figure 5a). In the eastern part of the study area where elevations are above 5,300 m a.s.l., they were less than 1,000°C-d, while in the western part with elevations lower than 5,300 m a.s.l., they varied from 1,000 to 3,510°C-d. The thawing indices in the outer part, especially in the western rims, were larger than the interior part. The pattern of freezing indices appeared contrary to
that of the thawing indices (Figure 5b). The freezing indices ranged from 34 to 4,148 °C·d. Elevations below 4,600 m a.s.l. in the western part had freezing indices less than 1,000 °C·d. The map of the ratio of thawing indices over freezing indices (Figure 5c) indicates...
considerably higher thawing indices in the western part, over four times higher than the freezing indices. The other areas, especially those with many parts higher than 5,000 m a.s.l. (colored in blue), had notably higher freezing indices, favoring permafrost formation. The areas with higher thawing indices are probably seasonally frozen ground with the assumption of no impacts from local factors.

4.2 | Soil clusters across the subregions

A total of five clusters were formed by the k-prototype approach for the soils across the subregions. As shown in Figure 6, clusters 1 and 4 are primarily found in subregions R2 and R1. Clusters 2 and 3 are mostly found in R3. Cluster 5 is scattered throughout the three subregions. The areal percentages of the clusters in each subregion are presented in Table 1. Clusters 4 (38.17%), 1 (45.07%), and 2 (38.54%) occupy the most areas in subregions R1, R2, and R3, respectively. Subregion R3 has the most diverse soil clusters. Clusters 1 (22.41%), 2 (25.22%), and 4 (22.9%) constitute a majority of the total for the three subregions. Cluster 5 occupies the least area (12.26%) of the total.

The primary characteristics of the soil clusters are summarized in Table 2. All the clusters have a dominant soil texture of sandy soil. Clusters 1 and 5 are characterized with a flat topography (slope <1.5°, TWI <0.21), relatively low elevation, and moderate precipitation (62–202 mm). Cluster 3 represents an extremely high alpine environment, which is typical in R3 that located in the central Gaize area, featuring high elevations from 5,122 m a.s.l. upwards, steepest relief (largest relief and slope in all clusters), and generally high precipitation. Clusters 2 and 4 are similar in elevation but distinct from each other in soil moisture conditions with regard to TWI and precipitation. Among all the factors, the topographical factors are shown to have the most pronounced effects on spatial partitioning.

4.3 | Optimal soil parameter values for soil clusters

The values of the soil parameter E were optimized for the soil clusters, as listed in Table 3. The values for clusters 1, 2, and 5 are >1, suggesting those soils do not favor the formation and preservation of permafrost. In contrast, the parameter values for soil clusters 3 and 4 are <1, indicating those soils favor permafrost development. In particular, cluster 3 has the lowest value (0.767) of the soil parameter among all clusters. Apart from the high elevations that cluster 3 represents, the local factors there such as high precipitation and soil parent material are beneficial to the formation of permafrost, in line with the interpretation for the lowest E value.

4.4 | Distribution of soil clusters over the study area

The distribution of soil clusters over the entire Gaize area was predicted, as shown in Figure 7. Soils in the western part of the study area consist primarily of clusters 4 and 1. Cluster 1 is also seen in the northeast part and usually occurs together with cluster 5. Cluster 3 is mainly located in high elevations in the central Gaize area. Cluster 2 is most likely to occur in the northeast and southern Gaize area.

The areal proportions of soil clusters in Gaize are presented in Table 1. Cluster 4 covers a maximal area of around 30.67%, followed by soil clusters 1 and 2, which share 22.81% and 21.95%, respectively. Cluster 3 occupies the least area (9.88%), whereas cluster 5 covers an extent of 14.69%. The constitution of soil clusters in the entire Gaize area resembles that in the three subregions. It signifies to some degree the representativeness of the selected subregions to the entire study area.

4.5 | Simulation and assessment in the subregions

The simulated permafrost distributions in the three subregions are broadly in good agreement with the survey-based subregion maps in terms of spatial pattern (Figure 8), with k equal to 0.45, 0.69, and 0.61, and OA equal to 0.66, 0.88, and 0.84 for R1, R2, and R3, respectively (Table 4). Permafrost areas are consistently located in the northern R1 and seasonally frozen ground covers the majority of the southern R1. Seasonally frozen ground in both maps dominates in subregion R2 except in a small area in the southeast where permafrost is present. By contrast, permafrost dominates R3 and seasonally

**FIGURE 6** Spatial distribution of soil clusters in the three subregions: R1 (a), R2 (b) and R3 (c) [Colour figure can be viewed at wileyonlinelibrary.com]
Frozen ground occupies much smaller portions such as in the southwest quadrant. The modeled permafrost extents are 0.32, 0.27, and $2.24 \times 10^3$ km$^2$ for R1, R2, and R3, respectively, close to but slightly lower than the extents on the survey-based subregion maps (0.35, 0.33, and $2.35 \times 10^3$ km$^2$).

A total of 22 boreholes have been drilled in the three subregions. Despite their use in the survey-based maps, the borehole records do not directly contribute to the simulation and therefore can be used for assessment of the simulation. The frozen ground types at 19 borehole sites are correctly identified in the simulation. Three sites in subregion R2 that were revealed as permafrost through boreholes are incorrectly identified in both the survey-based map and the simulation. The three boreholes were located on the northern slope (ZK34), hilltop (ZK35), and the southern slope (ZK36). ZK34 is about 2 km from ZK36. MAGT was approximately $-1.0$ °C in ZK34 and $-0.2$ °C in ZK36. Although permafrost can be seen on the shady slopes, sunny slopes of high elevations, and the cap of the hill, the portion of permafrost area is more likely to be small in a 1-km grid because the average elevation of the grid is low. Seasonally frozen ground is thus likely to occur as presented in the simulation.

Among the three subregions, R2 ($k$: 0.69; OA: 0.88) and R3 ($k$: 0.61; OA: 0.84) have the highest simulation accuracy and R1 ($k$: 0.45; OA: 0.84) has the lowest accuracy.

### Table 1: Areal percentages of soil clusters in the individual subregions, all subregions, and the entire Gaize area

| Subregion | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 | Cluster 5 |
|-----------|-----------|-----------|-----------|-----------|-----------|
| R1        | 20.33%    | 17.17%    | 3.83%     | 38.17%    | 20.50%    |
| R2        | 45.07%    | 3.71%     | 4.07%     | 35.07%    | 12.08%    |
| R3        | 10.73%    | 38.54%    | 27.35%    | 12.85%    | 10.53%    |
| R1–R3     | 22.41%    | 25.22%    | 17.21%    | 22.90%    | 12.26%    |
| Gaize     | 22.81%    | 21.95%    | 9.88%     | 30.67%    | 14.69%    |

### Table 2: Primary characteristics of the soil clusters. Elev.: elevation; TWI: topographic wetness index; Precip.: precipitation

| Cluster | Elev. (m) | Relief (m) | Aspect | Slope (°) | Soil texture | TWI | Precip. (mm) |
|---------|-----------|------------|--------|-----------|--------------|-----|--------------|
| 1       | 4,635–5,264 | 41–100     | E–W    | 0.6–1.2   | Sandy soil   | 0.21| 100–202      |
| 2       | 4,801–5,744 | 75–173     | S      | 1.5–4.6   | Sandy soil   | 0.17| 150–201      |
| 3       | 5,122–6,121 | 257–507    | E–W    | 4.5–13.7  | Sandy soil   | 0.26| 121–246      |
| 4       | 4,700–5,334 | 114–250    | S      | 2.3–6.2   | Sandy soil   | 0.45| 105–189      |
| 5       | 4,520–5,238 | 14–57      | N      | 0.2–1.5   | Sandy soil   | 0.14| 62–192       |

### Table 3: Optimized values of the soil parameter $E$ for the five soil clusters

| Soil cluster | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 | Cluster 5 |
|--------------|-----------|-----------|-----------|-----------|-----------|
| Optimal $E$  | 1.114     | 1.043     | 0.767     | 0.906     | 1.064     |

---

**Figure 7** Inferred distribution of soil clusters over the entire Gaize area. [Colour figure can be viewed at wileyonlinelibrary.com]
OA: 0.66) has the lowest. As indicated in Figure 8, R2 and R3 possess more boreholes than R1 (only one borehole), so the survey-based maps on R2 and R3 are likely to be of better quality than the map of R1. Generally speaking, as the subregion maps serve as the target of optimization, the quality of subregion maps will inevitably impact the simulation accuracy. This indicates the importance of using reliable subregion maps to achieve good simulation performance. However, for the case in subregion R1 that shows relatively low accuracy against the survey map, we cannot simply conclude that the performance in R1 was compromised because the survey-based map in R1 may be subject to large uncertainty. In our simulation, all three subregion maps rather than just the R1 map were used for global optimization. In this way, the adverse effect from a target map with large uncertainty can be alleviated. The use of multiple subregion permafrost maps appears to be beneficial for the simulation accuracy.

By visually examining the simulated and survey-based maps in the three subregions, some noticeable disagreements are apparent. For example, there is more seasonally frozen ground in the northeast R3 in the survey map (Figure 8, c2) than in the simulation (Figure 8, c1). The freezing indices in the difference areas were 1.6–2.0 times larger than the thawing indices, as indicated in blue in Figure 5(c), showing that climate was advantageous for permafrost formation and preservation. The local soil conditions, represented as cluster 2 with a parametric value of 1.043, did not favor the development of permafrost, offsetting climatic advantages in forming permafrost. The difference areas are located in a valley filled with alluvial fine sandy soils and have an average elevation of about 5,100 m a.s.l. The local climate, topography, and soil conditions are well exemplified in the two boreholes drilled nearby. The borehole records indicate that ground temperature at 10 m depth approached 0°C and the thickness of permafrost was about 30 m. Permafrost in those areas is of high-temperature type. It thus can be reasonably inferred that the difference areas are more likely to be underlain by high-temperature permafrost as in the simulation, rather than seasonally frozen ground as found on the survey-based map.
Simulation and assessment in the Gaize area

The permafrost area was modeled to be $20.86 \times 10^3$ km$^2$ in Gaize (Figure 9a) and accounts for 51% of the total area (Table 4). Seasonally frozen ground occupies $20.32 \times 10^3$ km$^2$ in the simulation and constitutes 49% of the total area. The simulated permafrost area is slightly smaller than the survey-based map (Figure 9b) by $-0.57 \times 10^3$ km$^2$ (Table 4). Both maps show very close similarity with respect to the spatial pattern, with a Kappa coefficient of 0.69 and an OA of 0.89.

The survey-based map of Gaize was independently made by an empirical approach and was not input to the simulation. Hence, the resulting high consistency between them firmly supports that the proposed approach is able to well reproduce both the broad and the local characteristics of permafrost distribution. The simulation shows that permafrost is mainly distributed in hilly and mountainous areas, while seasonally frozen ground is mostly distributed in lower plains. The continuity of permafrost decreases as elevation decreases.

Some discrepancies between the simulation and the survey-based map can be identified. In the southwest portion of the study area (marked by rectangles in Figure 9), there is less permafrost in the simulation than on the survey-based map. The remote sensing-derived thawing indices are 1.5–2.0 times larger than those in the difference area (Figure 5c). The local soil conditions are primarily of clusters 1 and 5, $E$ values of which are both $>1$, indicating benign soil conditions for developing seasonally frozen ground. Both climatic factors and local factors suggest the occurrence of seasonally frozen ground in this difference area, as presented on the simulation map.

The Zou map (Figure 9c) depicts broadly similar spatial patterns of permafrost distribution to those on the two other maps, with a lower agreement ($\kappa = 0.48$) than the simulation ($\kappa = 0.69$) against the survey-based map. While the Zou map and the simulation map used the same MODIS LST data, they differ in the model used and in processing of the satellite data. The FROSTNUM/COP approach used for generating the simulation map takes into account spatially thermal and moisture inhomogeneity in soils. As a result, some notable differences have seen in the southwest and northeast portions (marked by rectangles in Figure 9). The Zou map (Figure 9c) has most seasonally frozen ground distributed in the southwest portion among the three maps and the simulation (Figure 9a) has a distribution between the two. Given the extremely high elevations (>5,000 m a.s.l. on average) in this portion, complete coverage of seasonally frozen ground seems unrealistic. Likewise, in the northeast portion, both the ratios of freezing indices over thawing indices and the calibrated soil parameters probably rule out the existence of permafrost as shown in the Zou map. However, as indicated in Figure 9, we lack sufficient observations in those difference areas to truly support the results. The comparisons can only theoretically highlight the advantages of the proposed approach in accounting for both climate and local factors in determining the occurrence of permafrost. In addition, the proposed approach, like any other empirical/semi-physical model, assumes vertically uniform soil conditions and is limited in considering soil heterogeneity, the existence of ground ice, and ancient relict permafrost.

CONCLUSIONS

This study proposes an effective approach, namely FROSTNUM/COP, for mapping permafrost distribution with satellite data and subregion permafrost distribution maps in a large and data-scarce area. The approach uses spatial clustering and parameter optimization techniques to inversely compute the soil parameters, which have been introduced to the extended surface frost number model for characterizing spatial heterogeneity of soil conditions.

The approach is advanced in accounting for both climate and local factors in determining the occurrence of permafrost and its ability to produce a high-quality permafrost distribution map. Its effectiveness has been demonstrated in the experimental case of the Gaize area, where three subregion maps were developed as inputs and the simulation was validated against the survey-based Gaize map and a recently published map. The case study reveals a very close similarity between the simulation and survey-based map with regard to spatial pattern, as measured by a high Kappa coefficient (0.69) and overall accuracy (0.89). Analyses of the discrepancies between the
comparative maps show the simulation map is better than the survey-based map and the recently published map.

Although it has been well evaluated for performance in this study, simulation accuracy will be affected by the quality of subregion maps and their representativeness in the study area, which calls for further investigations into the error sources and influencing factors. This new approach can be applied and is an effective means to map accurate permafrost distribution in a large data-scare area if reliable subregion approaches can be applied and is an effective means to map accurate permafrost distribution in a large data-scare area if reliable subregion permafrost maps exist.

ACKNOWLEDGEMENTS
This work was supported by National Natural Science Foundation of China grants (Nos. 41671055, 41931180 and 41971074). We thank the Cryosphere Research Station on Qinghai–Xizang Plateau, Chinese Academy of Sciences, for providing the survey data in the Gaize area, and thank Dr Shugong Wang from NASA for help with the English text.

REFERENCES
1. Zimov SA, Schuur EA, Chapin FR. Permafrost and the global carbon budget. Science. 2006;312(5780):1612–1613.
2. Jin H, Wei Z, Wang S, et al. Assessment of frozen-ground conditions for engineering geology along the Qinghai-Tibet highway and railway, China. Eng Geol. 2008;101(3–4):96–109.
3. Biskaborn BK, Smith SL, Noetzel J, et al. Permafrost is warming at a global scale. Nat Commun. 2019;10(1):211–264.
4. Heginbottom JA. Permafrost mapping: a review. Progr Phys Geogr. 2002;26:623–642.
5. Zhao S, Nan Z, Huang Y, Zhao L. The application and evaluation of simple permafrost distribution models on the Qinghai-Tibet plateau: evaluation of permafrost distribution models on QTP. Permafrost Periglac Process. 2017;28(2):391–404.
6. Qin Y, Wu T, Zhao L, et al. Numerical modeling of the active layer thickness and permafrost thermal state across Qinghai-Tibetan plateau. J Geophys Res Atmos. 2017;122(21):11,604–11,620.
7. Wu X, Nan Z, Zhao S, Zhao L, Cheng G. Spatial modeling of permafrost distribution and properties on the Qinghai-Tibet plateau. Permafrost Periglac Process. 2018;29(2):86–99.
8. Li X, Cheng G. A GIS-aided response model of high-altitude permafrost to global change. Sci China Earth Sci. 1999;42(1):72–79.
9. Nan Z, Li S, Liu Y. Mean annual ground temperature distribution on the Tibetan plateau: permafrost distribution mapping and further application. J Glaciol Geocryol. 2002;24:142–148. (in Chinese)
10. Riseborough DW. The mean annual temperature at the top of permafrost, the TTOP model, and the effect of unfrozen water. Permafrost Periglac Process. 2002;13(2):137–143.
11. Nelson FE, Outcalt SI. A computational method for prediction and regionalization of permafrost. Arctic Antarct Alpine Res. 1987;19(3):279–288.
12. Boeckli L, Brenning A, Gruber S, Noetzel J. A statistical approach to modelling permafrost distribution in the European Alps or similar mountain ranges. Cryosphere. 2012;6(1):125–140.
13. Obu J, Westermann S, Bartsch A, et al. Northern hemisphere permafrost map based on TTOP modelling for 2000-2016 at 1 km² scale. Earth-Science Reviews. 2019;193:299-316.
14. Zou D, Zhao L, Sheng Y, et al. A new map of permafrost distribution on the Tibetan plateau. The Cryosphere. 2017;11(6):2527–2542.
15. Shi Y, Niu F, Yang C, Che T, Lin Z, Luo J. Permafrost Presence/-Absence Mapping of the Qinghai-Tibet Plateau Based on Multi-Source Remote Sensing Data. Remote Sensing. 2018;10(2):309–327.
16. Cheng G. Influences of local factors on permafrost occurrence and their implications for Qinghai-Xizang railway design. Sci China Earth Sci. 2004;47(8):704–709.
17. Fang X, Luo S, Lyu S, et al. A simulation and validation of CLM during freeze-thaw on the Tibetan plateau. Adv Meteorol. 2016;2016:1–15.
18. Xiao Y, Zhao L, Dai Y, Li R, Pang Q, Yao J. Representing permafrost properties in CoLM for the Qinghai-Xizang (Tibetan) plateau. Cold Reg Sci Technol. 2013;87:68–72.
19. Yin ZF, Ouyang H, Chen H. Simulating soil freezing and thawing of temperate desert ecosystem on the Qinghai-Tibet plateau. Procedia Environ Sci. 2010;2:476–485.
20. Zhou J, Kinzelbach W, Cheng G, Zhang W, He X, Ye B. Monitoring and modeling the influence of snow pack and organic soil on a permafrost active layer, Qinghai-Tibetan Plateau of China. Cold Reg Sci Technol. 2013;90–91:38–52.
21. van der Velde R, Su Z, Ek M, Rodell M, Ma Y. Influence of thermodynamic soil and vegetation parameterizations on the simulation of soil temperature states and surface fluxes by the Noah LSM over a Tibetan plateau site. Hydrol Earth Syst Sci. 2009;13:759–777.
22. Chen H, Nan Z, Zhao L, Ding Y, Chen J, Pang Q. Noah modelling of the permafrost distribution and characteristics in the West Kunlun area, Qinghai-Tibet plateau, China. Permafrost Periglac Process. 2015;26(2):160–174.
23. Gisnås K, Ettelmüller B, Lussana C, et al. Permafrost map for Norway, Sweden and Finland. Permafrost Periglac Process. 2017;28(2):359–378.
24. Nguyen T, Burn CR, King DJ, Smith SL. Estimating the extent of near-surface permafrost using remote sensing, Mackenzie Delta, Northwest Territories. Permafrost Periglac Process. 2009;20(2):141–153.
25. Westermann S, Østby TI, Gisnås K, Schuler TV, Ettelmüller B. A ground temperature map of the North Atlantic permafrost region based on remote sensing and reanalysis data. Cryosphere. 2015;9(3):1303–1319.
26. Langer M, Westermann S, Heikenfield M, Dorn W, Boije K. Satellite-based modeling of permafrost temperatures in a tundra lowland landscape. Remote Sens Environ. 2013;135:12–24.
27. Zhao L, Wu T, Xie C, et al. Support geoscience research, environmental management, and engineering construction with investigation and monitoring of permafrost in the Qinhai-Tibet plateau, China. Bull Chin Acad Sci. 2017;10:113–122. (in Chinese)
28. Yao T. The second tibetan plateau scientific expedition and research (STEP). Sci Bull. 2019;64(7):417–420. https://doi.org/10.1016/j.scib.2019.03.033
29. Streletskaïi D, Biskaborn BK, Smith SL, Noetzel J, Viera G, Schoeneich P. Strategy and implementation plan 2016–2020 for the global terrestrial network for permafrost (GTN-P). Washington D.C.: The George Washington University; 2017.
30. Deozdov DS, Malkova VG, Romanovsky VE, et al. Monitoring of permafrost in Russia. Russian database and the international GTN-P project. 68th Canadian Geotechnical Conference. Quebec, Canada. 2015.
31. Burgess MM, Smith SL, Brown J, Romanovsky V, Hinckel K. Global terrestrial network for permafrost (GTN-P): permafrost monitoring contributing to global climate observations. Geol Surv Canada Curr Res. 2000:46:1–8.
32. Nan Z, Li S, Cheng G, Huang P. Surface frost number model and its application to the Tibetan plateau. J Glaciol Geocryol. 2012;34(01):89–95. (in Chinese)
33. Luo D, Jin H, Marchenko SS, Romanovsky VE. Difference between near-surface air, land surface and ground surface temperatures and their influences on the frozen ground on the Qinghai-Tibet plateau. Geoderma. 2018;312:74–85.
34. Huang P, Nan Z. Estimation of 0-cm soil temperature over the Qinghai-Tibet plateau based on the wavelet analysis and adaptive network-fuzzy inference system. *J Glaciol Geocryol*. 2013;35:74-83. (in Chinese)

35. Yu W, Nan Z, Wang Z, Chen H, Wu T, Zhao L. An effective interpolation method for MODIS land surface temperature on the Qinghai-Tibet plateau. *IEEE J-STARS*. 2015;8(9):4539-4550.

36. Kondrashov D, Ghil M. Spatio-temporal filling of missing points in geophysical data sets. *Nonlinear Processes Geophys*. 2006;13:151-159.

37. Zhu A, Hudson B, Burt J, Lubich K, Simonson D. Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Sci Soc Am J*. 2001;5(65):1463-1472.

38. Ji J, Pang W, Zhou C, Han X, Wang Z. A fuzzy k-prototype clustering algorithm for mixed numeric and categorical data. *Knowl Base Syst*. 2012;30:129-135.

39. Poil R, Kennedy J, Blackwell T. Particle swarm optimization: an overview. *Swarm Intell J*. 2007;22(1):387-408.

40. Safavian SR, Landgrebe D. A survey of decision tree classifier methodology. *IEEE Trans Syst Man Cybern Syst*. 1991;21:660-674.

41. Qiao Y, Zhao L, Pang Q, Chen J, Zou D, Gao Z. Characteristics of permafrost in Gerze county on the Tibetan plateau. *J Glaciol Geocryol*. 2015;37(06):1453-1460. (in Chinese)

42. Chen J, Zhao L, Sheng Y, et al. Some characteristics of permafrost and its distribution in the Gaize area on the Qinghai-Tibet plateau, China. *Arctic Antarct Alpine Res*. 2016;48(2):395-409.

43. Shangguan W, Dai Y, Liu B, et al. A China data set of soil properties for land surface modeling. *J Adv Model Earth Syst*. 2013;5(2):212-224.

44. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37-46.

45. Congalton RG. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens Environ*. 1991;37:35-46.

46. Brennan RL, Prediger DJ. Coefficient kappa: some uses, misuses, and alternatives. *Educ Psychol Meas*. 1982;41:687-699.

47. Zhao L, Cheng G, Li S, Zhao X, Wang S. Thawing and freezing processes of active layer in Wudaoliang region of Tibetan plateau. *Chin Sci Bull*. 2000;45(23):2181-2187.

How to cite this article: Hu J, Zhao S, Nan Z, Wu X, Sun X, Cheng G. An effective approach for mapping permafrost in a large area using subregion maps and satellite data. *Permafrost and Periglac Process*. 2020;31:548–560. [https://doi.org/10.1002/ppp.2068](https://doi.org/10.1002/ppp.2068)