Permafrost Distribution along the Qinghai-Tibet Engineering Corridor, China Using High-Resolution Statistical Mapping and Modeling Integrated with Remote Sensing and GIS

Fujun Niu, Guoan Yin, Jing Luo, Zhanju Lin and Minghao Liu

1 State Key Laboratory of Frozen Soils Engineering, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China; luojinabcd@163.com (J.L); zhanjulin@lzb.ac.cn (Z.L); liuminghao@lzb.ac.cn (M.L.)

2 South China Institute of Geotechnical Engineering, School of Civil Engineering and Transportation, South China University of Technology, Guangzhou 510006, China

* Correspondence: niufujun@lzb.ac.cn (F.N.); yinguoan@lzb.ac.cn (G.Y.)

Received: 18 October 2017; Accepted: 30 January 2018; Published: 1 February 2018

Abstract: Permafrost distribution in the Qinghai-Tibet Engineering Corridor (QTEC) is of growing interest due to the increase in infrastructure development in this remote area. Empirical models of mountain permafrost distribution have been established based on field sampled data, as a tool for regional-scale assessments of its distribution. This kind of model approach has never been applied for a large portion of this engineering corridor. In the present study, this methodology is applied to map permafrost distribution throughout the QTEC. After spatial modelling of the mean annual air temperature distribution from MODIS-LST and DEM, using high-resolution satellite image to interpret land surface type, a permafrost probability index was obtained. The evaluation results indicate that the model has an acceptable performance. Conditions highly favorable to permafrost presence (≥70%) are predicted for 60.3% of the study area, declaring a discontinuous permafrost distribution in the QTEC. This map is useful for the infrastructure development along the QTEC. In the future, local ground-truth observations will be required to confirm permafrost presence in favorable areas and to monitor permafrost evolution under the influence of climate change.

Keywords: statistical model; permafrost map; remote sensing; Qinghai-Tibet Plateau

1. Introduction

The IPCC Fifth Assessment Report [1] indicates that most permafrost has been degrading since the last little ice age and the rate has increased recently, as evidenced by permafrost temperature increasing and a positive trend of the active layer thickness. Permafrost degradation has an impact on surface and subsurface hydrologic conditions, soil strength properties, and ecosystems [2]. Mountain permafrost was widely considered to possibly influence slope stability [3] and hydrological systems [4] and poses a challenge to economic development in high mountain areas. Permafrost in China was mainly found on the Qinghai-Tibet Plateau (QTP), the largest lower latitude permafrost region (1.05 × 10^6 km^2) in the world. Warming of permafrost during the last decades was well documented along the Qinghai-Tibet Engineering Corridor (QTEC) [5–8]. Changes in the active layer thickness and permafrost temperature due to climate warming and surface disturbances have major ecological and engineering implications [8–12]. The thawing of ice-rich permafrost has an impact on the stability of the slopes and the thermokarst [13–15], which could lead to the instability of the Qinghai-Tibet Railway (QTR) [16]. Climatic change and its associated effects on ground surface evolution, landscape dynamics, and natural hazards make it important to map permafrost distribution...
on the QTP. Therefore, improved methods for mapping permafrost distribution are essential to designing road and pipelines and to understanding the dynamics of alpine ecosystems.

The construction of a new expressway from Golmud to Lhasa has been proposed along the QTEC. Since the permafrost condition at a site would affect the engineering design and cost of road construction, a detailed and present-day knowledge of permafrost distribution and its relationships with geomorphology were essential in the corridor. Present available permafrost maps were mainly at low resolutions, which varied from scales of 1: 600,000 to 1: 10,000,000, although many permafrost maps have been compiled since the early 1960s [17]. At more local scales, factors that affect local microclimate and surface energy balance e.g., slope, aspect, local hydrology, vegetation cover, geology, and snow cover strongly influenced the permafrost features [18,19]. A good understanding of these effects on permafrost occurrence was significant, as it may provide some hints on the techniques and measures we can use to artificially simulate similar effects [20–22]. Thus, a detailed and up-to-date permafrost map of the QTEC was required for mitigating potential engineering problems associated with permafrost-affected terrain. It will also help to identify areas for further investigation so that permafrost areas can be avoided or, if necessary, engineering solutions can be designed to maintain the physical and thermal state of permafrost.

Many models already exist for estimating the spatial distribution of permafrost in regions of the European Alps [23–25], and the Arctic [26,27]. These models may be of the equilibrium, empirical-statistical, or process-based types and have been widely used at regional and local scales [3,28]. Permafrost mapping based on geophysical techniques or process-based types was expensive, time consuming, and spatially restrictive due to the difficult detection and monitoring of permafrost on the QTP, although these techniques can provide a detailed and robust transient thermal state of permafrost. Furthermore, the lack of sufficient and reliable data for calibration and validation probably was one of the most important limitations for permafrost modeling [25]. Empirical-statistical models describing the distribution of mountain permafrost based on geomorphological permafrost indicators and topographic and climatic predictors were a simple yet effective approach toward a first assessment of its distribution at a regional scale [29–32]. Furthermore, remote sensing as a permafrost monitoring tool was under continuous development with fine spatial and temporal resolution data from satellites such as Landsat-8, SPOT-5, and GF-1 and 2 (Gaofen-1, 2 satellite). This technique has the potential to provide a valuable and cost-effective mean for mapping and monitoring near-surface permafrost conditions, as well as seasonally frozen ground [33]. High-resolution image and elevation data acquired by satellite can be used to interpret the geomorphological features such as slope failure, thermokarst [16,34], and biophysical features such as vegetation, topography, and surface hydrology [26,35]. A number of studies have demonstrated the usefulness of this approach for permafrost mapping [29,36–39].

Therefore, the target of this study was to map the potential permafrost distribution in the QTEC (91°E–95°E, 32°N–36°N) based on a logistic regression model linking the permafrost existence probability to surface variables that included vegetation type and climatic conditions at a high resolution of 30 m. For this purpose, we investigated the surface characteristics (e.g., vegetation, soil) and permafrost conditions along the QTEC based on the permafrost survey position and boreholes that were carried out when the QTR and Qinghai-Tibet highway (QTH) were built. Considering the potential incoming solar radiation, vegetation type, and the mean annual air temperature as potential predictors, they were compiled and developed from remote sensing data.

2. Study Area

The study area was located in the central QTP and encompassed a 550 km long and 40 km wide section (22,351 km²) of the QTP (Figure 1a,b), extending from Xidatan to Anduo, which were the northern and southern boundaries of permafrost occurrence (Figure 1b), respectively, bounded by latitude 32°–36°N and longitude 91°–95°E (Figure 1c). This corridor is likely to become a locus of many developmental projects because constructions of a gas pipeline, QTP, QTR, and electric
Remote Sens. 2018, 10, x FOR PEER REVIEW 3 of 19
literature indicated that near surface deposits were dominated by coarse materials such as gravel and sandy soils [10,40]. These specific geomorphologic and sedimentary patterns resulted in significant differentiations in permafrost features [10]. Climatically, it was located in an extremely continental climate zone, favoring clear skies and high solar radiation. The mean annual air temperatures (MAAT) were commonly between $-6.5\, ^\circ\text{C}$ and $-2.0\, ^\circ\text{C}$, with the annual total precipitation ranging from 250 mm to 450 mm, which occurred mostly as rainfall between May and August [6,8,40]. The majority of the plateau had a free snow cover in winter [40]. Vegetation type was characterized as alpine meadow and steppe with the coverage ranging from 0.3 to 0.9. The detailed information of the vegetation type was discussed in Section 3.3.

Figure 1. Study area map. (a) The Qinghai-Tibetan Plateau location in China; (b) map of Permafrost on the Qinghai-Tibetan Plateau (Modified based on the permafrost map by Li and Cheng [41]; (c) topographic map of the whole study area based on a 30-m DEM data, including mainly local geographic names of mountains and upland basins.

3. Methods and Materials

In adopting the approach to permafrost probability modeling, we collected and compiled the necessary data with field work and remote sensing. In order to create predictor variables such as mean annual air temperature (MAAT), potential incoming solar radiation (PISR), and vegetation type, MODIS-LST products, a 30 m-DEM (ASTER GDEM), and high resolution (2 m) GF-1 images were used. The indicator variable representing permafrost presence and absence was based on the thermal
boreholes. The field point measurements of climate and surface features were used for calibration and validation.

3.1. Field Data

We carried out the field work from August, 2014 to October, 2015 along the QTEC, mainly near the QTR and QTH due to the difficult assessment. The sample locations were mainly based on two sources: (1) thermal boreholes along the QTH and QTR [6,7,40], and (2) geological survey boreholes by the Road Survey and Design Institution during the period August, 2015 to October, 2015 (15 m depth). These boreholes could record the permafrost condition (presence or absence) clearly. At each location, we recorded the soil type, surface features, vegetation type present, and measured the topography (aspect, slope) with a geological compass. A hand-held GPS (Trimble, Sunnyvale, CA, USA) was used to measure the geographical coordinate of each location. Finally, we got 344 samples in different vegetation, topographic, and geologic settings (Table 1). Information on permafrost was recorded based on existing thermal boreholes in the study area. In general, permafrost was found in all vegetation types with different frequencies. Vegetation where permafrost was found with intermediate to high frequency included degrading alpine meadow (82.5%) and alpine steppe (91.7%). Statistics of field data showed that vegetation classes were strongly related to near-surface soil type; however, this relationship needed more detailed ground investigation for identifying. Active-layer thickness (ALT) in different vegetation classes had great variations. For example, in alpine meadow the ALT ranged from 0.9 m to 5.3 m. The ALT varied from 1.1 m to 4.1 m for degrading alpine meadow. In the alpine steppe, the range was from 1.0 m to 6.5 m. For the sparse grassland, the ALT was between 0.8 m and 7.5 m. Generally, the ALT was thinnest in the alpine meadow and degrading alpine meadow. In addition, the permafrost temperature (mean annual ground temperature at 10 m depth, MAGT) varied in different vegetation classes. For most sites in alpine meadow and degrading alpine meadow, the MAGT was colder than that in other vegetation classes.

Table 1. Summary of field data collected in different vegetation, topographic, and geologic settings.

<table>
<thead>
<tr>
<th>Vegetation Type</th>
<th>Near-Surface (0–50 cm) Soil Type</th>
<th>Permafrost Present/% Absent</th>
<th>Active Layer Thickness (m)</th>
<th>Permafrost Temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpine meadow</td>
<td>Sandy loam</td>
<td>85/82.5</td>
<td>0.9–5.3</td>
<td>−3.9–0</td>
</tr>
<tr>
<td>Degrading alpine meadow</td>
<td>Sandy loam</td>
<td>44/77.2</td>
<td>1.1–4.1</td>
<td>−3.6–0</td>
</tr>
<tr>
<td>Alpine steppe</td>
<td>Sand with gravel</td>
<td>33/91.7</td>
<td>1.0–6.5</td>
<td>−2.9–0</td>
</tr>
<tr>
<td>Sparse grassland</td>
<td>Sand with gravel</td>
<td>37/46.8</td>
<td>0.8–7.5</td>
<td>−2.6–0</td>
</tr>
<tr>
<td>Bare ground</td>
<td>Stone and gravel</td>
<td>37/53.6</td>
<td>1.2–6.2</td>
<td>−3.1–0</td>
</tr>
</tbody>
</table>

Seven climate stations were available in this study (Figure 1c, Table 2). They were located in different terrains and recorded the near ground surface (1.5 m above the ground surface) air temperature every 4 h. The field air temperature was used to verify the relationship between MODIS-LST and air temperature.

Table 2. Locations of the weather stations and number of years with observations.

<table>
<thead>
<tr>
<th>Weather Station Name</th>
<th>Longitude (° N)</th>
<th>Latitude (°E)</th>
<th>Elevation (m)</th>
<th>Averaged MAAT</th>
<th>Record Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>K 966 *</td>
<td>94.04</td>
<td>35.72</td>
<td>4587</td>
<td>−4.4</td>
<td>2014–2016</td>
</tr>
<tr>
<td>K 980</td>
<td>94.05</td>
<td>35.61</td>
<td>4736</td>
<td>−5.8</td>
<td>2014–2016</td>
</tr>
<tr>
<td>Wudaoling</td>
<td>93.08</td>
<td>35.19</td>
<td>4612</td>
<td>−4.5</td>
<td>2012–2016</td>
</tr>
<tr>
<td>K 3035</td>
<td>92.98</td>
<td>34.99</td>
<td>4593</td>
<td>−3.5</td>
<td>2014–2016</td>
</tr>
<tr>
<td>Beiluhe</td>
<td>93.93</td>
<td>34.82</td>
<td>4648</td>
<td>−3.4</td>
<td>2005–2016</td>
</tr>
<tr>
<td>Tuotuohe</td>
<td>92.43</td>
<td>34.22</td>
<td>4533</td>
<td>−4.0</td>
<td>2005–2016</td>
</tr>
<tr>
<td>Anduo</td>
<td>91.72</td>
<td>32.40</td>
<td>4780</td>
<td>−3.3</td>
<td>2014–2016</td>
</tr>
</tbody>
</table>

* K 966, K 980—Mileage of the QTR, K 3035- Mileage of the QTH.
3.2. Statistical Model and Validation

Logistic regression model has been widely used in spatial permafrost modeling [31,32,39]. For the detailed theory of this statistical model, one can refer to the above literature. This approach estimates the probability of permafrost presence \((Y = 1)\) or absence \((Y = 0)\) based on a set of log odds of dependent variable. It solves this problem (Equation (1)):

\[
\ln \left( \frac{P(Y = 1)}{1 - P(Y = 1)} \right) = B_0 + B_1 x_1 + B_2 x_2 + \ldots + B_n x_n
\]

in which \(P(Y = 1)\) is the probability of permafrost existence, \(B_0\) the intercept, and \(B_n\) the repression coefficient of the explanatory variable \(x_n\). Therefore, a non-linear transformation from the linear repression creates an S-shaped distribution function. The predictors related to the probability of permafrost presence are solved by a multivariate global logistic repression model (Equation (2)):

\[
P(Y = 1) = \frac{1}{1 + e^{-(B_0 + B_1 x_1 + B_2 x_2 + \ldots + B_n x_n)}}
\]

in which \(e\) is the base of the natural logarithm.

The above process step was undertaken using the statistical program package SPSS 20 (IBM Corp., 2011, Armonk, NYC, USA). To avoid a chance occurrence of anomalous model performance, we ran the logistic regression model ten times, each time with approximately two-thirds of the randomly selected sampled data (Monte Carlo random number generation in MATLAB) to train the model and the remaining one-third to validate the model. We used the coefficients, which were produced from the run with highest classification accuracy for testing data, as input variable classes. Additionally, model performance was evaluated using the area under the receiver operating characteristic curve (AUROC) in software program ROC-AUC [42]. In this study, the curve showed the probability of detecting observed permafrost occurrences and absences for the whole range of possible decision thresholds that could be used to dichotomize predicted odds into permafrost presence/absence. To simulate the final accuracy, we compared the model results with spatial field measurements from three typical study areas—Xidatan area, Beiluhe basin, and Anduo area, in which permafrost was investigated in detail in recent studies (Figures S1 and S2).

3.3. Remote Sensing-Independent Variables

3.3.1. Ground Surface Type

Mountain permafrost environments were characterized by a strong heterogeneity on small spatial scales, so that high resolution (HR) remote sensors were generally required. The imagery of GF-1 was useful to extrapolate the surface features due to its high spatial resolution (2.0 m × 2.0 m) and wide area coverage (>60 km × 60 km). GF-1 data has five spectral bands: panchromatic (0.45–0.90 μm), blue (0.45–0.52 μm), green (0.52–0.59 μm), red (0.63–0.69 μm), and NIR (0.77–0.89 μm). The image was preprocessed by China Centre for Resources Satellite Data and Application (CCRSDA) before we used in this study. The process included radiometric calibration, atmospheric correction, geometric correction, and image sharpening. On average, the orthorectification process achieved a root mean square error (RMSE) of less than 4 pixels in both x and y directions. Finally, 45 images were acquired for the study area from the CCRSDA. These images covered the period from July 2014 to October 2014, and were selected based on the time with little or no cloud. Meanwhile, this period was at two different phenological stages: the start of vegetation growing season and the late summer. The ground surface type classes were defined mostly based on plant community features, as well as their GF-1 spectral signature (Table 3). The object-oriented classification and visual interpretation methods were used to generate maps of land cover. Moreover, the object-oriented image feature extraction method was suitable for the HR data. This process step was completed using the ENVI
software (http://www.harrisgeospatial.com/docs/ExtractFeatures.html) under the tool of ENVI Feature Extraction. This tool provided a guided workflow with two main parts: finding the object and extracting the features. Table 3 summarized the object features and extraction rules adopted in this study. Normalized difference vegetation index (NDVI) was one of the most widely used indices applied to mapping vegetation. The NDVI of different land cover types had obvious difference. Analysis of image texture (from band 4) showed the textural differences between different ground surface types. Additionally, different landforms have different shape and colour characteristics within the satellite images. Finally, nine land-cover classes were derived from the GF-1 images with an output resolution of 30 m. Theoretically, classification accuracy should be assessed against existing maps that were established using other methods; however, for our study area, such maps were not available. In order to estimate the classification accuracy, we directly investigated and compared the classified image with the field conditions at 400 GPS locations.

3.3.2. Mean Annual Air Temperature (MAAT)

To build a high-resolution response variable representing air temperature conditions, the MODIS-LST products (1 km resolution) were used in this study to produce near surface air temperature dataset. MODIS L3 data were available from NASA Warehouse Inventory Search Tool website. The used LST (land surface temperature) products contain day-and night-time surface temperature based on the generalized split window approach [43]. The assumption that the arithmetic average of the maximum and minimum LSTs can represent the daily mean LST has been well established [44,45]. A time series of daily LST average was compiled using at least one daytime and one nighttime data, and then we calculated the mean annual LST. The details of processing were similar to Zou et al. [46]. Finally, we get an annual time series of LST from 2012 to 2016.

Previous research showed that MODIS-LST did not closely match the real ground surface temperature, but was close to air temperature in permafrost environment [47]. On the QTP, similar rule was also found due to the mixed pixel covering different vegetation type and water body [48,49]. We compared the mean annual MODIS-LST with field point measurements of air temperature. According to this comparison, we assumed that the MODIS-LST could present the MAAT spatially. The statistical-empirical model was based on an equilibrium assumption that meant we should use temperature averages over long term rather than a single year. Therefore, 5-year averaged MAAT was used in this study.

High-resolution spatially distributed information on MAAT was modeled for period 2012–2016 by a simple model (Equation (3)) from MODIS-LST products. Daily temperature data from available MODIS-LST products was used to (i) calculate the appropriate lapse rate by the Equation (3) and a 1 km-DEM; and (ii) lapse to 30 m grid point with a 30 m-DEM by Equation (3), given that the lapse rate ($\Delta T/\Delta H$) in a 1 km grid only moderately varied.

$$Lapse\ rate = \frac{\Delta T}{\Delta H} = \frac{MAAT_i - T_o}{H_i - H_o}$$

In which $MAAT_i$ is the temperature of each grid ($^\circ$C), $T_o$ the averaged value of MODIS-LST products ($^\circ$C), $H_i$ the elevation of each grid (m), and $H_o$ the averaged elevation (m).

3.3.3. Solar Radiation (PISR)

Spatially distributed solar radiation (Wh/m²) was calculated using “Area Solar Radiation” tool and the ASTER GDEM data with a 30 × 30 m resolution in ArcGIS 9.3 (ESRI, 2009, Redlands, CA, USA). Annual PISR was calculated at a 10-day interval with a half hour time resolution between 4 a.m. and 10 p.m., for the whole year, since the snow cover was free during winter. Default values (such as diffuse proportion 0.3, transmissivity 0.5) representing general clear skies were used as radiation parameters due to lack of field meteorological data.
Table 3. Object features of NDVI, texture, shape, and colour for the satellite image used in this study.

<table>
<thead>
<tr>
<th>Class ID</th>
<th>Object</th>
<th>Definition</th>
<th>NDVI</th>
<th>Texture</th>
<th>Shape</th>
<th>Image Colour</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alpine meadow</td>
<td>Vegetation cover &gt;0.7, dominated by <em>Kobrcia parva</em></td>
<td>0.4–1.0</td>
<td>Uniform and clear</td>
<td>Continuous distribution and boundary</td>
<td>Red, light red, and pink</td>
</tr>
<tr>
<td>2</td>
<td>Degrading alpine meadow</td>
<td>Vegetation cover between 0.3 and 0.8, dominated by <em>Kobrcia parva</em> and <em>Stipa purpurea</em></td>
<td>0.2–0.6</td>
<td>Fine and smooth</td>
<td>Irregularity</td>
<td>Red, light red, and pink</td>
</tr>
<tr>
<td>3</td>
<td>Alpine steppe</td>
<td>Sand and gravel land covered by vegetation (cover &lt;0.25), such as <em>Carex moorcroftii</em></td>
<td>0.15–0.3</td>
<td>Fine</td>
<td>Irregularity</td>
<td>Red, light red, and pink</td>
</tr>
<tr>
<td>4</td>
<td>Sparse grassland</td>
<td>Bare land with very sparses vegetation (cover b0.2), such as <em>Stipa purpurea</em></td>
<td>0.15–0.25</td>
<td>clear</td>
<td>Irregularity</td>
<td>Pink and light red</td>
</tr>
<tr>
<td>5</td>
<td>Bare ground</td>
<td>Exposed lands without vegetation cover, such as bedrock outcrops with sand and gravel, roads, and work-yards.</td>
<td>0–0.15</td>
<td>Fine and uniform</td>
<td>Irregularity</td>
<td>off white</td>
</tr>
<tr>
<td>6</td>
<td>Lake</td>
<td>Deep water, including thermokarst lakes</td>
<td>&lt;0</td>
<td>Fine, clear, and uniform</td>
<td>Obvious geometrical characteristic, such as oval</td>
<td>Blue, dark blue, and light blue</td>
</tr>
<tr>
<td>7</td>
<td>River</td>
<td>Shallow water including streams</td>
<td>&lt;0</td>
<td>Fine and clear</td>
<td>banded</td>
<td>Blue, dark blue, and light blue</td>
</tr>
<tr>
<td>8</td>
<td>Bench land</td>
<td>Fluvial outwash</td>
<td>0–0.1</td>
<td>Fine and uniform</td>
<td>zonal or sheet</td>
<td>off white</td>
</tr>
<tr>
<td>9</td>
<td>Snow/glacier</td>
<td>Mainly distributed on the high mountain peak.</td>
<td>&lt;0</td>
<td>Fine and homogeneous</td>
<td>Distributed along contour line</td>
<td>White and uniform</td>
</tr>
</tbody>
</table>
4. Results

4.1. Remote Sensing Inputs

Main remote sensing products were developed: land cover map from GF-1 images (Figure 2a), MAAT modeled from MODIS-LST product, and 30-m DEM (Figure 2b), and PISR from a 30-m DEM (Figure 2c). The land cover map (Figure 2a) showed that the land cover could be classified into 9 types, and the vegetation (4 types) covered the most area (over 85%) (Figure 3). According to the land cover map, the vegetation mainly distributed below 5100–5400 m a.s.l. on the south-facing slopes and 4600–5100 m a.s.l. on the north-facing slopes. Alpine meadow dominated the whole study area and occupied 25% of the QTEC. There were approximately 34,937 lakes interpreted on the images over the study area. Lake areas ranged from 17.5 m² to 111.2 km² with mean area of 11,429 m². Snow cover was distributed over the high mountains such as Tanggula mountain and Kunlunshan. This was limited snow cover distribution, which occupied only 1.4% of the total area (Figure 3). Generally, snow cover distributed above 5100–5600 m a.s.l. on the south-facing slopes and 4900–5200 m a.s.l. on the north-facing slopes. Compared with 400 field-sampled locations, overall accuracy of the land-cover classification was 92.6% (Table 4). Frequency statistics of spatially-distributed annual PISR were shown in Figure 4c. The PISR had a normal distribution with a mean value of $2.1 \times 10^6$ Wh/m² and a standard deviation of $1.46 \times 10^5$ Wh/m². The spatially-distributed MAAT was calculated from MODIS-LST products. Figure 4b indicated that downscaled MODIS-LST (modeled MAAT, Equation (3)) highly agreed with the measured MAAT ($R^2 = 0.87$). The mean modeled MAAT was $-3.9 \, ^\circ\text{C}$ (with standard deviation (SD) of 0.8 °C), which was similar to the observed mean MAAT of 4.1 °C with a range of $-3.3$~$-5.8$ °C (Table 2). The mean error (ME) of the modeled MAAT was $-0.42$ °C with a root mean square error (RMSE) of 0.57 °C. The result indicated a 5-year mean MAAT of 0 °C isotherm altitude at 4250 m a.s.l. in Xidatan area, which rose southward to 4700 m a.s.l. in Anduo area.

Figure 2. Remote sensing-independent variable products. (a) Land cover type map of the study area-QTEC, derived from GF-1 multispectral data using object-oriented classification approach. (b) Modeled mean annual air temperature distribution in the QTEC. (c) Spatially distributed solar radiation.
4.2. Statistical Permafrost Model

Statistical analysis was based on a dataset of 344 field samples. The minimum and maximum classification accuracies achieved with the training data ranged from 83.0% to 87.4%, whereas those for testing ranged from 80.7% to 87.7%, assuming the common cutoff value of 0.5. The average overall classification accuracies achieved for ten runs were 85.4% and 83.5% for training and testing data, respectively (Table 5). For all ten runs, the significant level was well above 0.05, suggesting the model adequately fit the input data. We selected the coefficients as input variables, which were produced from regression fits, as well as all coefficient estimates, was highly significant (< 0.001, Table 5). As expected, sampling locations were randomly selected from across the study area.

<table>
<thead>
<tr>
<th>Class ID</th>
<th>Class Name</th>
<th>Number of Sites</th>
<th>User’s Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alpine meadow</td>
<td>103</td>
<td>92.1</td>
</tr>
<tr>
<td>2</td>
<td>Degrading alpine meadow</td>
<td>57</td>
<td>88.7</td>
</tr>
<tr>
<td>3</td>
<td>Alpine steppe</td>
<td>36</td>
<td>85.5</td>
</tr>
<tr>
<td>4</td>
<td>Sparse grassland</td>
<td>79</td>
<td>91.3</td>
</tr>
<tr>
<td>5</td>
<td>Bare ground</td>
<td>69</td>
<td>87.2</td>
</tr>
<tr>
<td>6</td>
<td>Lake</td>
<td>20</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>River</td>
<td>10</td>
<td>98</td>
</tr>
<tr>
<td>8</td>
<td>Bench land</td>
<td>16</td>
<td>91</td>
</tr>
<tr>
<td>9</td>
<td>Snow/glacier</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td><strong>Average user’s accuracy</strong></td>
<td></td>
<td>92.6</td>
</tr>
</tbody>
</table>

The regression model's overall classification accuracy was 92.6% (Table 4). Frequency statistics of spatially-distributed potential incoming solar radiation in the QTEC (Y = modeled MAAT, X = measured MAAT) were shown in Figure 4. The result indicated a 5-year mean MAAT of 0 °C isotherm altitude at 4250 m a.s.l. in Xidatan area, which rose southward to 4700 m a.s.l. in Anduo area.

\[
\text{MAAT} = 3.085 + 0.835 \times \text{PISR} - 0.486 \times \text{MAAT} - 1.188
\]

\[
\text{MAAT} \text{ (°C)} = 0.42 \pm 0.835 \times \text{PISR} - 0.486 \times \text{MAAT} - 1.188 \quad \text{(Equation (3))}
\]

\[
\text{RMSE} = 0.57 \quad \text{(°C)}
\]

The result was similar to the observed mean MAAT of 4.1 °C with a range of −5.8 °C to 3.3 °C (Table 2). The mean error (ME) of the modeled MAAT was −0.486 ± 0.072 °C, and the coefficient of determination (R²) was 0.978. The spatially-distributed MAAT was calculated from MODIS-LST products. Figure 4b indicated that downscaled MODIS-LST (modeled MAAT, Equation (3)) highly correlated with measured MAAT (Equation (2)). Significant correlation was found between MAAT and solar radiation. The relationship between measured MAAT and modeled MAAT was linear, with a correlation coefficient of 0.87 (Figure 4c). The spatially-distributed solar radiation had a normal distribution with a mean value of 2.1 × 10⁶ Wh/m2 and a standard deviation of 1.46 × 10⁵ Wh/m2. The spatially-distributed potential incoming solar radiation was normally distributed with a mean value of 2.1 × 10⁶ Wh/m2 and a standard deviation of 0.8 °C, which was similar to the observed mean MAAT of 4.1 °C with a range of −5.8 °C to 3.3 °C (Table 2). The mean error (ME) of the modeled MAAT was −0.486 ± 0.072 °C, and the coefficient of determination (R²) was 0.978. The spatially-distributed solar radiation had a normal distribution with a mean value of 2.1 × 10⁶ Wh/m2 and a standard deviation of 1.46 × 10⁵ Wh/m2. The spatially-distributed solar radiation was calculated from MODIS-LST products.
the model run with highest classification accuracy for testing data (Table 5). The regression model’s overall fit, as well as all coefficient estimates, was highly significant (bootstrap \( p < 0.001 \), Table 5). As expected, both MAAT and solar radiation were negatively related to permafrost occurrence. A probability threshold of 0.7 was chosen to discriminate between “permafrost probable \(( p \geq 70\% )\)” and “permafrost improbable \(( p < 70\% )\)” locations, and thus for delineating the potential permafrost extent. This classification cutoff was based on the optimal cutoff of 0.73, which was calculated by ROC-AUC program.

### Table 5. Statistical coefficients and accuracy assessments for the model.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients</th>
<th>Bias</th>
<th>SE</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vegetation</td>
<td>−1.058</td>
<td>0.041</td>
<td>1.14</td>
<td>0.000</td>
</tr>
<tr>
<td>PISR</td>
<td>−1.188</td>
<td>−0.835</td>
<td>5.79</td>
<td>0.000</td>
</tr>
<tr>
<td>MAAT</td>
<td>−0.486</td>
<td>−0.072</td>
<td>0.313</td>
<td>0.000</td>
</tr>
<tr>
<td>Constant</td>
<td>3.085</td>
<td>0.825</td>
<td>4.943</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### VALIDATION

<table>
<thead>
<tr>
<th>Run</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highest *</td>
<td>83.0%</td>
<td>87.7%</td>
</tr>
<tr>
<td>Average</td>
<td>85.4%</td>
<td>83.5%</td>
</tr>
</tbody>
</table>

Note: * Highest classification accuracy for testing data among ten runs; ** Ten-runs’ average classification accuracy for training the model.

#### 4.3. Spatial Permafrost Distribution

Spatially-distributed probabilities of permafrost presence condition were computed in ArcGIS with raster calculator, inserting the regression’s coefficient estimates in Equation (2) (Table 5). The spatial distribution of the permafrost probability in the QTEC was depicted in Figure 5a. In general, the potential permafrost areas tended to decrease in upland plain in the central QTEC (Tuotuohe area and Kaixinling area). More favorable conditions were concentrated in the higher areas of the study areas, such as Kunlunshan (>4900 m a.s.l.), Fenghuoshan (>4800 m a.s.l.), and Tanggula mountain (>5000 m a.s.l.). In contrast, lower scores \(( p < 50\% )\) were associated with lower hillslopes, valley bottoms, and upland plains.

Considering a \( p > 70\% \) and excluding rivers, lakes, and glacier surfaces (total 452 km\(^2\)), highly favorable conditions for permafrost occurrence were inferred for 60.3% of the study area, or 13,196 km\(^2\), indicating the discontinuous permafrost distribution (underlying 50–90% of the landscape) in the QTEC (Table 6). If considering a general score >50%, the potential permafrost would cover 87.8% of the study area, or 19,236 km\(^2\). In the north-boundary of the QTEC—Xidatan area, permafrost was probable \(( p > 0.7 )\) above 4350 m a.s.l. on north-facing slopes and above 4700 m a.s.l. on south-facing slopes. In the central ranges of the QTEC, the average lower boundary of permafrost occurrence lay at approximately 4700 m a.s.l. In the south boundary of the QTEC-Anduo areas, the calculated limits were higher; permafrost was probable above 4700 m a.s.l. on south-facing slopes and above 4900 m a.s.l. on north-facing slopes. Upland plains were usually above 4500 m a.s.l. (Xieshuihe, Beiluhe, Chumaerhe, and Tuotuohe); however, there seemed no obvious lower limit of permafrost distribution.

Comparisons of modeled permafrost distribution with independent ground-truth observations were performed in this study. We compared the results with the observations in three areas—Xidatan, Beiluhe, and Anduo (Figure 5b–d), although not all of the research data have been published (completed and under submitting, Figures S1 and S2). Table 7 showed that the overall consistency between modeled permafrost existence and direct permafrost observations was 75.0%, considering that Xidatan and Anduo were the boundaries of permafrost distribution on the QTP, where permafrost was warm, and this accuracy was acceptable.
Table 6. Distribution of areas favorable for permafrost in the QTEC.

<table>
<thead>
<tr>
<th>Permafrost Probability (%)</th>
<th>Area * (km²)</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–20</td>
<td>402</td>
<td>1.8</td>
</tr>
<tr>
<td>20–50</td>
<td>2260</td>
<td>10.3</td>
</tr>
<tr>
<td>50–70</td>
<td>6040</td>
<td>27.6</td>
</tr>
<tr>
<td>70–90</td>
<td>10,204</td>
<td>46.6</td>
</tr>
<tr>
<td>90–100</td>
<td>2992</td>
<td>13.7</td>
</tr>
<tr>
<td>Total</td>
<td>21,899</td>
<td>100</td>
</tr>
</tbody>
</table>

* Note: Rivers, lakes, and glacier surfaces (452 km²) excluded.

Table 7. Comparison of direct permafrost observations.

<table>
<thead>
<tr>
<th>Observed and Other Studies</th>
<th>Predicted Permafrost Existence</th>
<th>Consistency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Xidatan (n * = 13)</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Beiluhe (n = 20)</td>
<td>14</td>
<td>1</td>
</tr>
<tr>
<td>Anduo (n = 3)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Overall (n = 36)</td>
<td>23</td>
<td>4</td>
</tr>
</tbody>
</table>

Note: * number of the boreholes.

Figure 5. Modeled permafrost probability distribution in the all QTEC excluding rivers, lakes, and glacier surfaces (a), the Xidatan area (North boundary) (b), Beiluhe basin (central area) (c), and Anduo area (South boundary) (d).
5. Discussion

5.1. Predictor Variables for the Model Inputs

Land cover, e.g., vegetation type, has large effects on the development and distribution of permafrost in the corridor, because of vegetation’s influence on surface energy balances and associations with soil characteristics [50], consistent with field observations in this study (Table 1). Shallow lithology is an important factor influencing permafrost temperatures and extents because of varying thermal offsets in different materials in the corridor [51]. While there are surficial geology investigations in this study, they had poor spatial accuracies and only empirical-statistical relations between shallow soil and vegetation type were demonstrated. Therefore, geological factors were not considered in this research due to unavailable geology maps. Its effects should be quantitatively studied in the further studies. The significant effects of potential incoming solar radiation on mountain permafrost extents have been identified in various alpine areas [52–54], consistent with model estimates in this study.

Climatic indices such as MAAT were the primary drivers of landscape-level permafrost extent [54]. Commonly, MAAT is a relatively available variable. Spatial extrapolation of MAAT with the choice of appropriated lapse rates in rugged terrain is very important for permafrost distribution modeling. It is well known that near-surface lapse rates vary in time and space, even over short distances [55–57]. Therefore, there are still large uncertainties from use of a spatially uniform lapse rate for temperature extrapolations [31]. Meanwhile, measurements of near-surface lapse rates on the QTP are rare, impeded by low density of long-term climate stations. It is significant to employ a simple but efficient model to calculate spatial air temperature distribution over large area [32]. In this study, the high-resolution near-surface air temperature distribution was modeled from MODIS-LST products and a 30 m DEM by a simple model (Equation (3)). The overall temperature lapse rate obtained in this study was $-0.45 \degree C$ per 100 m with a standard deviation of 0.42 $\degree C$, which was lower than the average temperature decrease ($-0.6 \degree C$ per 100 m) in the free atmosphere [38], but reasonably close to the average lapse rate of $-0.5 \degree C$ per 100 m on the QTP [57]. The precision of the modeled MAAT ($R^2 = 0.87$, $ME = -0.42 \degree C$, $RMSE = 0.57 \degree C$, $SD = 0.8 \degree C$) announces the uncertainties in combining different resolution datasets. The main issues for the model here are that there are still landform variations, even over short distance within a 1 km grid. In addition, the quality of DEM products can also affect the calculation accuracy. Considering the effects of different underlying surface on MODIS-LST, overall modeled MAAT may have a warm bias, even though they are compared with the limited availability of field air temperature data (Figure 4b). However, this level of uncertainty indicates that our research approach is likely applicable in order to obtain high-resolution near-surface air temperature distribution on the QTP. Given the low density of available climate stations along the QTEC for model inputs and verifications, more high elevation stations are urgently needed in the further studies.

5.2. Permafrost Model Performance

Sophisticated numerical models can be applied to address large-scale climate impact on permafrost distribution for lowland areas [59]. For more local applications in mountain areas, permafrost indictors such as temperature measurements under a stable snow cover (BTS) or landforms (rock glaciers) may be more suitable [60], due to lack of physical parameters that characterize permafrost and a mass of computing resource need of numerical models. The study presented here is a first-time attempt to simulate the permafrost distribution using a multi-dimensional statistical relationships between permafrost existence and meteorological-land cover-derived in the QTEC. A major drawback for the data basis is the lack of data in steeper terrains (i.e., higher mountains such as Tanggula mountain areas). Figure 6 shows the model variability and uncertainty for three sections of the study area. In Xidatan section, MAAT and surface conditions show little variations, and the degree of solar radiation and local hydrology has a large effect on permafrost existence and model output. Unlike upland plain areas such
as Beiluhe basin, vegetation cover and type show a negative relation to permafrost existence in this study. This is probably due to the low number of quantitative data points and limited consideration for shallow lithology and hydrological influences. Variations in specific local environments and processes in upland plain areas such as permafrost distribution depicted in Yin et al. [61] cause the most model variability and uncertainty in our study. For Anduo section, there are only three field sites; however, the MAAT and PISR are the main factors governing the permafrost existent and model result due to the site location (south boundary of permafrost distribution) on the QTP.

This type of approach used in our study is well known and widely used in mountain areas such as the Alps, and the rock glacier as the empirical foundation for permafrost favorability models is often used to validate the results. However, there are not enough studies referring to rock glaciers on the QTP. It is difficult for researchers to make appropriate assumptions for permafrost existence. Evidence of direct thermal borehole is more suitable for our study, at least they were placed representatively according to a meaningful sampling designing.

In general, our study gives a coherent picture of permafrost distribution, but does not reproduce active-layer thickness and permafrost temperature. Therefore, this approach is not useful for evaluating the impact of climate change on permafrost condition. However, this map is very useful for current infrastructure construction, such as engineering problem solutions associated with permafrost and further investigation for permafrost areas.

5.3. Comparison with Previous Studies

The most widely used permafrost distribution map on the QTP was compiled by Li and Cheng [41] (Figure 1b). This map synthesized field data, literature, aerial photographs, satellite images, and many relevant maps. Wang et al. applied a 0.5 °C MAGT isotherm to depict the permafrost distribution on the QTP [62]. This threshold was interpolated based on the relationship between elevation/latitude and the MAGT observations along the QTH [63]. However, limited numbers of field observations were used for assessment of these maps, without independent validation [63]. The permafrost boundary was mainly based on a threshold of air temperature isotherms and modified in several regions.
using field data based on the authors’ knowledge. The approach presented in this study differed from these mainly in the algorithm employed to estimate permafrost probability integrated with high-resolution remote sensing data and in the complete statistical assessment of the model results. A direct comparison of our results (areas with score $\geq 70\%$) to these maps showed that our results could provide more detailed permafrost distribution at a local scale, and permafrost is discontinued in the QTEC. Furthermore, model-model comparisons were performed for assessing uncertainties in permafrost distribution. The permafrost zonation index (PZI) of Gruber [64] was an independent global-scale empirical modelling effort at a 1km resolution based on downscaled reanalysis data. For comparison, our result was resampled to PZI resolution. Permafrost areas with a PZI $\geq 0.7$ (12,044 km$^2$, excluding lake, river, and glacier surface) were smaller than the areas in this study (13,196 km$^2$) along the QTEC. This study had an AUROC of 0.73, which is similar to the AUROC value obtained in recent permafrost modelling studies in the Alps [25] and Andes [32].

Overall, compared to previous empirically based studies [31,32,41,62,65,66], the permafrost modelling approach used in this work integrates remote sensing products with more flexible, nonlinear models of potential permafrost distribution in comparison with previous research approaches.

5.4. Permafrost Temperature and Effects of Climate Change

Climate warming has resulted in significant degradation of permafrost [1]. The global average increase in MAAT has been 0.3–0.6 °C since 1880, and on the QTP MAAT has increased 0.3–0.5 °C during the past 40 years [67]. Permafrost near the lower limit of permafrost distribution can be more sensitive to degradation processes due to the possible effects of climate change [68]. Over the past 50 years, and particularly since the 1980s, air temperature and precipitation have distinctly increased in the central QTEC [61], indicating that over the last decades there has been warming and wetting. Along the QTEC, ALT was reported to be increasing at 6.3–7.8 cm·yr$^{-1}$ over a period from 1995 to 2010 [7,8], which was higher than the rates of 1.33 cm·yr$^{-1}$ for the period 1981–2010 outside the QTEC [69]. The average permafrost temperature rise at the depth of 6.0 m was 0.02 °C·yr$^{-1}$ over a period from 2006 through 2010 [8]. These warmings could lead to geotechnical problems and potential hazards to important roads or infrastructures in the QTEC (e.g., QTH and QTR) [14,70]. In this study, the permafrost probability map can serve as an up-to-date resource to assess permafrost conditions and uncertainties in mountain research and practical applications such as infrastructure planning.

6. Conclusions

In this study, a statistical permafrost distribution model provided an insight into mountain permafrost distribution in the Qinghai-Tibet Engineering Corridor, China. We derived the high-resolution, spatially permafrost probability map by developing logistic regression models from field data and remote sensing data, and integrating the model in a GIS framework. From this study the following conclusions can be drawn:

1. The results of the land cover type and MAAT model emphasized the importance of the GF-1 satellite image to derive ground surface parameters, and the appropriateness of the MODIS-LST product to determine temperature distribution with scarce and heterogeneous temperature records from weather stations. Moreover, the results of the MAAT model can provide valuable data to study other climatically driven earth surface features widely and relatively easily on the QTP.

2. The approach used here has given a valid picture of permafrost distribution at a local scale. The map accuracies were assessed using independent field observations. The model results suggested that permafrost was discontinuous and occupied about 60.3% of the QTEC, excluding rivers, lakes, and glacier surfaces. Lower limits of permafrost occurrence ($p > 70\%$) increased from 4350 m a.s.l. in the north to 4700 m a.s.l. in the south.
3. Finally, the findings of this research contributed mainly to improving the general knowledge about permafrost distribution in the QTEC, providing valuable information to government for infrastructure planning, such as pipeline or railroad routes. The high-resolution vegetation map will also serve as a baseline map to identify areas of permafrost degradation and ecosystem changing under the background of climate warming. Additional research, in particular, taking into account ground-truthing and local conditions, is necessary in order to refine the present model and evaluate possible permafrost change.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/10/2/215/s1.

Figure S1. Photo of the thermal borehole distribution for permafrost investigation in Xidatan. The results have been submitted to Journal of Mountain Science by Luo et al. (2017). (Jing Luo, Fujun Niu, Zhanju Lin, Guoan Yin, Minghao Liu, variations in the northern permafrost limit in the Qinghai-Tibet Plateau over the last four decades, Journal of Mountain Science, 2017, under revising); Figure S2. Photo of the thermal borehole distribution for permafrost investigation in Anduo. The results have not been submitted.

Acknowledgments: This research was supported financially by the State Key Program of National Natural Science of China (Grant No. 41730640), the National Natural Science Foundation of China (Grant No. 41601069; Grant No. 41571065), Independent Project of the State Key Laboratory of Frozen Soils Engineering (Grant No. SKLFSE-ZQ-46), and the Science Technology Research and Development Plan of China Railway General Corporation (Grant No. 2016G003-D). Changes made upon constructive comments by three anonymous reviewers and an academic editor further improved the paper.

Author Contributions: F.N. wrote the paper, analysed the field data as well as derived the methods. G.Y. analysed the remote sensing data, overviewed the whole process, reviewed the paper, and provided comments. J.L., reviewed the paper, provided comments, and provided the field data of Xidatan and Anduo used as a reference for assessing the accuracy. Z.L. and M.L. assisted in the field work and interpretation of the comparison results.

Conflicts of Interest: The authors declare no conflict of interest.

References


47. Hachem, S.; Duguay, C.; Allard, M. Comparison of MODIS-derived land surface temperatures with ground surface and air temperature measurements in continuous permafrost terrain. *Cryosphere* 2012, 6, 51. [CrossRef]


64. Gruber, S. Derivation and analysis of a high-resolution estimate of global permafrost zonation. Cryosphere 2012, 6, 221. [CrossRef]


© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).