PIC: Comprehensive R package for permafrost indices computing with daily weather observations and atmospheric forcing over the Qinghai–Tibet Plateau

Lihui Luo¹, Zhongqiong Zhang¹, Wei Ma¹, Shuhua Yi², Yanli Zhuang³

¹State Key Laboratory of Frozen Soils Engineering, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu Province 730000, China
²State Key Laboratory of Cryospheric Sciences, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu Province 730000, China
³Linze Inland River Basin Research Station, Key Laboratory of Inland River Basin Ecohydrology, Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu Province 730000, China

Correspondence to: Lihui Luo (luolh@lzb.ac.cn)

Abstract. An R package permafrost indices computing (PIC) was developed, which integrates meteorological observations, remote sensing data, and field measurements to compute the factors or indices of permafrost and seasonal frozen soil. At present, 16 temperature/depth-related indices are integrated into the R package PIC to estimate the possible change trends of frozen soil in the Qinghai–Tibet Plateau (QTP). These indices include the mean annual air temperature, mean annual ground surface temperature, mean annual ground temperature, seasonal thawing/freezing n factor (nt/nf), thawing/freezing degree-days of air and ground surface (DDTa/DDTs/DDFa/DDFs), temperature at the top of the permafrost, active layer thickness, and maximum seasonal freeze depth. The PIC package supports two computational modes, namely, the stations and region calculation that enables statistical analysis and intuitive visualization on the time series and spatial simulations. Over 10 statistical methods were adopted to evaluate these indices in stations, and a sequential Mann-Kendall trend test and spatial trend method were adopted. Multiple visual manners display the temporal and spatial variabilities on the stations and region. The data sets of 52 weather stations and a central region of QTP were prepared and simulated to evaluate the temporal–spatial change trends of permafrost with the climate. Simulation results show extensive permafrost degradation in QTP, and the temporal–spatial trends of the permafrost conditions in QTP were consistent with those of previous studies. The PIC package
will serve engineering applications and can be used to assess the impact of climate change on permafrost.

1 Introduction

Permafrost, which is soil, rock, or sediment with temperatures that have remained at or below 0 °C for at least two years, is a key component of the cryosphere. The upper layer in permafrost regions is called the active layer, which undergoes seasonal freezing and thawing. Below this layer lies permafrost, the upper surface of which is called the upper limit of permafrost or permafrost table. Changes in permafrost can affect water and heat exchanges, carbon budgets, and natural hazards with the climate change. Permafrost occurs mostly in high latitudes and altitudes with long, cold winters and thin winter snow cover (e.g., the Arctic, Antarctica, Alaska, Northern Russia, Northern Canada, Northern Mongolia, and the Qinghai–Tibet Plateau (QTP)) (Riseborough et al., 2008; Yi et al., 2014b; Zhang et al., 2008a). Over half of the QTP land is underlain by permafrost (Ran et al., 2012). The temperature in QTP has increased by 0.25 °C per decade over the past 50 years (Li et al., 2010; Liu et al., 2006; Shen et al., 2015; Yao et al., 2007). Climate-induced warming of the near-surface atmospheric layer and a corresponding increase in ground temperatures will lead to substantial changes in the water and energy balance of regions underlain by permafrost (Hilbich et al., 2008). Such an increase can warm the ground through energy exchange at the surface and result in significant permafrost degradation. The distribution and changes of permafrost with climate is necessary for infrastructure development, ecological and environmental assessments, and climate system modeling (Luo et al., 2017; Luo et al., 2012; Zhang et al., 2014).

Given the possibility of future climate warming, an evaluation of the magnitude of changes in the ground thermal regime has become desirable to assess the possible eco-environmental responses and their impact on the infrastructure in QTP. Permafrost modeling maximizes quantitative methods, such as analytical, numerical, or empirical methods, to predict the thermal condition of the ground in environments where permafrost may be present (Harris et al., 2009; Lewkowicz and Bonnaventure, 2008; Riseborough, 2011; Riseborough et al., 2008; Yi et al., 2014a; Zhang et al., 2008b). At present, dozens of different factors or indices are used to evaluate the characteristics and dynamics of permafrost presence or absence (Riseborough, 2011; Riseborough et al., 2008), including freezing/thawing index, mean annual air temperature (MAAT), mean annual ground
temperature (MAGT), mean annual ground surface temperature (MAGST), temperature at the top of permafrost (TTOP), and active layer thickness (ALT), among others. Thereafter, the type and distribution of frozen soil can be classified in a variety of manners depending on the index sizes and magnitude. For example, frozen soil can be divided into highly stable, stable, substable, transitional, unstable, and extremely unstable permafrost, as well as seasonal frozen soil that depends on the size of MAGT (Chen et al., 2012; Ran et al., 2012). These indices can be used to evaluate and predict the temporal and spatial variation in the thermal response of permafrost to the changing climate conditions and properties of Earth’s surface and subsurface properties in one, two, or three dimensions (Juliussen and Humlum, 2007; Nelson et al., 1997; Riseborough et al., 2008; Wu et al., 2010; Zhang et al., 2005). Accordingly, successfully summarizing and categorizing a variety of frozen soil indices require permafrost modeling that concerns analytical, numerical, and empirical methodologies to compute the past and present condition of permafrost. The Stefan solution (Nelson et al., 1997), Kudryavtsev’s approach (Kudryavtsev et al., 1977), TTOP model (Smith and Riseborough, 1996), and Geophysical Institute Permafrost Lab model (Romanovsky and Osterkamp, 1997; Sazonova and Romanovsky, 2003) are several important developments for permafrost modeling in recent years. Permafrost is a subsurface feature that is difficult to directly observe and map. These methods integrate the effects of air and ground temperatures, topography, vegetation, and soil properties to map permafrost spatially and explicitly (Gisnás et al., 2013; Jafarov et al., 2012; Zhang et al., 2014). Weather observation data, including air and soil temperatures with different depths, are the main inputs for single-point simulation. These indices consist mainly of temperature-related and depth-related indices. The temperature-related indices depict the status of air or land surface temperature in frozen soil environments, whereas the depth-related indices reveal the status of the active layer. Preparing atmospheric forcing data sets, snow depth and density, vegetation types, and soil classes are generally required for multi-dimensional simulation, which came from multi-source data fusion, particularly remote sensing and ground observation data.

The current lack of open source software on permafrost modeling over QTP is a problem. Although many scientists in China have field data and models on hand, the integration of data and models into a new open source model can facilitate the deepening of the discussion and unfolding of permafrost research on QTP. Given the current condition of permafrost modeling in QTP, a comprehensive R package permafrost indices computing (PIC) was developed to compute permafrost and seasonal frozen soil indices. The goal is to determine the solutions to maintain or build the engineers in a manner that provides guidance...
for the future of highway and high-speed railway design and construction in QTP, as well as further understand the effects of climate change on the permafrost dynamics over QTP. Therefore, the proposed software integrates meteorological observations, remote sensing data, field measurements, and model simulations.

The remainder of this paper is organized as follows. Section 2 describes the prepared data sets, methodology of permafrost modeling, and statistical methods for stations and region. Section 3 presents a detailed description of the functions provided by PIC and the workflow. Section 4 demonstrates the application of the proposed software for the stations and region. Section 5 discusses several benefits and limitations of PIC. Lastly, Section 6 presents the conclusions.

2 Data and Methods

2.1 Data and parameters

**Daily weather observations.** Meteorological data were obtained from the China Meteorological Administration (CMA, http://www.cma.gov.cn/), particularly from permanent meteorological stations across QTP. A total of 52 weather stations with daily meteorological records (i.e., from 1951 to 2010) were selected, including the daily mean, maximum (max) and minimum (min) air temperatures, wind speed, observed and corrected precipitation, evaporation, air humidity, atmospheric pressure, sunshine duration, daily mean, max and min ground surface temperatures, and soil temperature with different depths (i.e., 5, 10, 15, 20, 40, 50, 80, 160, and 320 cm). These data have been corrected under specification for surface meteorological observation and quality control of CMA.

**Atmospheric forcing data set.** The QT Engineering Corridor (QTEC), which is located at the center of QTP, was selected in preparing the atmospheric forcing data. Global Land Data Assimilation System (GLDAS, https://ldas.gsfc.nasa.gov) and the weather station data of the surrounding QTEC were merged to produce a new data set for 1980 to 2010 with a daily 0.1° temporal–spatial resolution (Rui and Beaudoin, 2011).

**Parameters.** The parameters for the ground conditions were prepared based on vegetation and soil classification (Bicheron et al., 2008; Nachtergaele et al., 2009), field observations, and topographic maps. The parameter data have two data sets: one for weather stations and another for the QTEC region. Table 1 and Figure 1 show the detailed information of the data and
parameters.

2.2 Permafrost modeling

The PIC package enables the calculation of MAAT, MAGST, MAGT, seasonal thawing/freezing $n$ factor ($n_t/n_f$), thawing/freezing degree-days of air and ground surface ($DDT_a/DDT_s/DDF_a/DDF_s$), TTOP, ALT, and maximum seasonal freeze depth (FD). These permafrost and seasonal frozen soil indices that employ the following function were illustrated.

$A_s$ is the annual temperature amplitude at the ground surface, where $T_{\text{max}}$ and $T_{\text{min}}$ are the annual max and min temperatures, respectively, at the ground surface. $A_s$ can be calculated as follows:

$$A_s = T_{\text{max}} - T_{\text{min}}$$  \hspace{1cm} (1)

$L$ is the volumetric latent heat of fusion, $\rho$ is the dry density of soil, and $W$ is the water content of the soil in percentages.

$$L = \frac{800 \times \rho \times W}{100}$$  \hspace{1cm} (2)

$DDT_a$ and $DDT_s$ are the thawing degree-days of the air and ground surface (Celsius degree-days), respectively. $DDF_a$ and $DDF_s$ are the freezing degree-days of the air and ground surface (Celsius degree-days), respectively, where $T_a$ and $T_s$ are the air and ground temperatures, respectively, and $n$ is the number of days in a year (Juliussen and Humlum, 2007).

$$DDT_a = \sum^n T_a, \quad T_a > 0$$  \hspace{1cm} (3)

$$DDF_a = \sum^n T_a, \quad T_a < 0$$  \hspace{1cm} (4)

$$DDT_s = \sum^n T_s, \quad T_s > 0$$  \hspace{1cm} (5)

$$DDF_s = \sum^n T_s, \quad T_s < 0$$  \hspace{1cm} (6)

$P$ is assigned a value of 365 days. MAAT and MAGST can be computed as follows:

$$\text{MAAT} = \frac{DDT_a - DDF_a}{P}$$ \hspace{1cm} (7)

$$\text{MAGST} = \frac{DDT_s - DDF_s}{P}$$ \hspace{1cm} (8)

MAGT is the soil temperature in (Wu and Zhang, 2010). $T_z$ is the ground temperature at any time $t$ and depth $z$ below a ground surface. MAGT can be computed (Juliussen and Humlum, 2007; Riseborough et al., 2008) as follows:

$$T_{z,t} = T_a + A_s \times e^{-2z \sqrt{\pi/\alpha P}} \times \sin\left(\frac{2\pi t}{P} - z \times \sqrt{\pi/\alpha P}\right)$$ \hspace{1cm} (9)
MAGT = $\overline{T_{x,t}}$, $z \cong 15$ & $t = 86400$.                       \hspace{1cm} (10)

$n_t$ and $n_f$ can be computed (Riseborough et al., 2008) as follows:

$$n_t = \frac{DDT_a}{DDT_\Delta}$$ \hspace{1cm} (11)

$$n_f = \frac{DDF_a}{DDF_\Delta}.$$ \hspace{1cm} (12)

Two methods serve the same purpose when computing TTOP and ALT. The subscripts $S$ and $K$ stand for the Smith and Kudryavtsev functions (Kudryavtsev et al., 1977; Smith and Riseborough, 1996), respectively.

$$TTOP_S = \frac{n_t \times \lambda_t \times DDT_a - n_f \times \lambda_f \times DDT_s}{\lambda_f \times \lambda_p} \times \frac{\text{MAGST}}{\lambda_s} \times \frac{\text{MAGST}}{\lambda_s} \times \frac{\pi}{A_s} \times \frac{1}{A_s^2}.$$ \hspace{1cm} (13)

$$TTOP_K = \frac{\lambda^*}{\lambda^*} \times \left\{ \begin{array}{ll}
\lambda_f, & \text{if numerator < 0} \\
\lambda_t, & \text{if numerator > 0}
\end{array} \right.$$ \hspace{1cm} (14)

$$\lambda^* = \left\{ \begin{array}{ll}
\lambda_f, & \text{if numerator < 0} \\
\lambda_t, & \text{if numerator > 0}
\end{array} \right.$$ \hspace{1cm} (15)

The maximum thawing depth or ALT uses the Stefan and Kudryavtsev functions (Kudryavtsev et al., 1977; Riseborough et al., 2008), where $L$ is the latent heat of fusion of ice ($3.34 \times 10^5$ J/kg).

$$\text{Alt}_S = \sqrt{\frac{2 \times A_s \times DDT_a}{L \times \rho \times (W - W_u)}}.$$ \hspace{1cm} (16)

$$A_s = \frac{A_s - T_s}{100} - \frac{L}{2 \times C_T}.$$ \hspace{1cm} (17)

$$Z_c = \frac{2 \times (A_s - T_s) \times (x + l) \times P_{SN} \times C_T}{2 \times \pi}.$$ \hspace{1cm} (18)

$$\text{Alt}_K = \sqrt{\frac{2 \times A_s \times TTOP_K \times (x + l) \times P_{SN} \times C_T}{2 \times \pi}}.$$ \hspace{1cm} (19)

Freeze_depth$_S$ is the maximum seasonal freezing depth that uses the Stefan function, which can be computed as follows:

$$\text{Freeze_depth}_S = \sqrt{\frac{2 \times A_s \times DDT_a}{L \times \rho \times (W - W_u)}}.$$ \hspace{1cm} (20)
2.3 Statistical methods

Statistical analysis can facilitate the evaluation of the change trend and overall performance of the model simulation. In particular, each statistic has strengths and weaknesses; thus, we adopted over 10 statistical methods to evaluate these indices in station computing for time series data. The quantitative statistics include the slope, $y$-intercept, Pearson’s correlation coefficient ($R$), coefficient of determination ($R^2$), root mean square error (RMSE), standard deviation (SD), ratio of scatter (RS), normalized RMSE (NRMSE), Nash–Sutcliffe efficiency (NSE), RMSE-observations standard deviation ratio (RSR), percent bias (PBIAS), normalized average error (NAE), variance ratio (VR), and index of agreement ($D$) (Jafarov et al., 2012; Legates and McCabe, 1999). The sequential Mann-Kendall (MK) trend test was used to statistically assess whether there was a shift in trends of the climate factors and permafrost indices (Fraile, 1993). The original MK trend test can be calculated as follows:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sign}(x_j - x_i), \quad (i = 2, 3, 4 ... n)$$ (21)

$$\text{sign}(x_j - x_i) = \begin{cases} 1 & \text{if } x_j - x_i > 0 \\ 0 & \text{if } x_j - x_i = 0 \\ -1 & \text{if } x_j - x_i < 0 \end{cases}$$ (22)

Two sequential series $u_i$ value can be calculated as follows:

$$u_i = \frac{S_i - E(S_i)}{\sqrt{\text{var}(S_j)}}, \quad (i = 1, 2, 3 ... n)$$ (23)

The two series for MK trend test, a progressive one and a backward one, were sets up. If they cross each other and diverge beyond a specific threshold value and exceeding the confidence level of 95%, then there is a statistically significant trend shift point.

The spatial trend can also be calculated through the function below. The index represents one permafrost index, $n$ represents the sequential years, and $\text{index}_i$ is the index values in year $i$. Taking ALT as an example, a positive trend value means that ALT was increasing, thereby indicating that permafrost degradation has intensified; a negative value means that ALT was decreasing, thereby indicating that permafrost degradation has a certain inhibition; and a zero trend suggests a lack of change (Chen et al., 2014; Stow et al., 2003).

$$\text{Trend} = \frac{n \times \sum_{i=1}^{n} \sum_{i=1}^{n} \text{index}_i \sum_{i=1}^{n} \text{index}_i - \sum_{i=1}^{n} \text{index}_i \sum_{i=1}^{n} \text{index}_i}{n \times \sum_{i=1}^{n} \text{index}_i \sum_{i=1}^{n} \text{index}_i}$$ (24)
3 Package description

PIC was developed in the R Language and Environment for Statistical Computing v. 3.3.3 and is distributed as open source software under the GNU-GPL 3.0 License. Therefore, the PIC code can be modified as required to meet the needs of every user. The source code can be downloaded at the GitHub repository (https://github.com/iffylaw/PIC). The R package PIC provides all the necessary functionality to perform the calculation, statistics, and drawing of permafrost indices with over 38 functions based on the user’s specific requirements (see Figure 2). The following packages are required in setting up the PIC (type library(PIC)): ggplot2 (Wickham et al., 2009), ggmap (Kahle and Wickham, 2013), RNetCDF (Michna and Woods, 2013); (Zambrano-Bigiarini and Rojas, 2013), and animation (Xie, 2013). These packages are automatically added to the users’ R library during installation. A data set that contains the daily weather observations, parameters, and information (i.e., from 1951 to 2010) of 52 weather stations in QTP were bundled into this package. However, the region data with the NetCDF format was placed in the GitHub repository. The data set variables excluded in the calculation can be used as reference or provide support to further develop PIC. These variables include wind speed, precipitation, evaporation, humidity, and soil temperature at different depths. PIC was primarily designed to compute indices of permafrost and seasonal frozen soil from observations and forcing data. Therefore, the current stable version of the program (v 1.0) includes functionalities that cover temperature-related indices (i.e., MAAT, MAGST, and TTOP) and depth-related indices (i.e., ALT and FD) that are commonly used in permafrost research.

The PIC package supports two computational modes: the station and region calculations that enable statistical analysis and visual displays on the time series and spatial simulations. The regional calculation adopts GIS approaches to compute each spatial grid. PIC was initially developed to address an immediate need for a reliable and easy-to-use program to estimate the temporal–spatial changes in frozen soil in QTP. Thus, the workflow comprises deliberately simplified steps involved throughout the entire process. Once PIC is installed, the workflow of the weather observations is considerably straightforward: (1) an index of a weather station for one year or multiple years is calculated, (2) an index of 52 weather stations from 1951 to 2010 is calculated, and (3) an index of all stations or permafrost stations from 1951 to 2010 is drawn through curve and spatial visualization. Step (1) is an optional step. The workflow of the forcing data has only two steps: (1) a total of 4 indices from
1980 to 2010 are calculated, including MAAT, $DDT_a$, $DDF_a$, and ALT and (2) the spatial statistic and visualization of these 4 indices are drawn. Table 2 describes most of these functions.

4 Examples

Several examples of the PIC use and application were presented here. This section highlights several significant features of the package in terms of specific functions, including station and region calculation, statistics, and visualization. However, PIC includes numerous illustrations from literature and possible detailed analyses. PIC has built-in station data. The data set comprises two tables (data frame), namely, QTP_ATM for daily weather observations and Station_Info for information and parameters in each station. Users can modify or adjust the parameters in the Station_Info and can use the data and parameters. Additional examples can be referenced in the GitHub repository.

4.1 Station calculation

We can use different functions in the R console to perform the calculations based on the selected method. For example, if a user wants to obtain a MAAT value for a certain year of a station, then he/she can enter the following command. TempName and data are optional in the MAAT function.

```
MAAT (Year = 1980, TempName = “Temperature”, data = QTP_ATM, SID = 52908)
```

User can also obtain the MAAT values for a specified period of years in a station.

```
MAAT (Year = 1980:2010, TempName = “Temperature,” data = QTP_ATM, SID = 52908)
```

The “Year” option can be assigned to a number and sequence. The other temperature/depth-related indices can also use the two inputs for the “Year” option. A user can obtain the values of all stations for an index. The “VarName” option can be equal to the function name in the Com_Indices_QTP function. The results can also be saved to a CSV file with column/row names. The case of the input VarName is supported.

```
Com_Indices_QTP (VarName = “MAAT”)
```

Given that the freezing/thawing index can be divided into freezing/thawing degree-days of the air and ground surface, the
VarName option should add "_air" or "_ground" at the end of the Freezing_index and Thawing_index. However, the abbreviation can also be utilized as the option input. The “Thawing_index_air” and “ta” are the same.

```plaintext
Com_Indices_QTP (VarName = “Thawing_index_air”)
Com_Indices_QTP (VarName = “ta”)
```

After the TTOP indices are computed, the stations that may have permafrost should be determined. The Exist_Permafrost function can determine and map the stations where permafrost exists. The probability of permafrost occurrence and most likely permafrost conditions are determined from the computing results of the Exist_Permafrost function (see Figure 3).

```plaintext
TTOP_S_QTP ≤ Com_Indices_QTP (VarName = “TTOP_Smith”)
TTOP_K_QTP ≤ Com_Indices_QTP (VarName = “TTOP_Kudryavtsev”)
Exist_Permafrost (plot = “yes”)
```

The QTP measurements have constantly been difficult. The data set has several null and anomalous values, as well as lead to a few anomalous values in computing the indices. Accordingly, these outlier values should be processed. The Outlier_Process function seeks the outlier values and sets these values to null values thereafter, which is an option because abnormal values have been processed in the Com_Indices_QTP.

```plaintext
Outlier_Process (MAAT_QTP[,1:stations])
```

### 4.2 Region calculation

A total of four indices can be computed with the atmospheric forcing data set in the PIC package. This package supports the NetCDF format data; thus, the package reads and writes a NetCDF file to support region computing. The input NetCDF files require a few forcing and parameter data. After the calculations, a user can compute the spatial statistics and draw the index changes through a GIF animation (see Sections 4.3 and 4.4).

```plaintext
Spatial_Pic (NetCDFName = “PIC_indices.nc”, StartYear = 1980, EndYear = 2010)
```

### 4.3 Statistics

The stat function contains all statistical methods for station calculation. The PIC package provides two statistical calculations
to compute the statistical value of all stations using Com_Stats_QTP: (1) the indices that vary with the change in the year and (2) the comparison of the same two indices for different computational methods. Options ind1 and ind2 were used; however, ind2 can be disregarded when computing the statistical values between single data and years.

\[
\text{Com}_\text{Stats}_\text{QTP} (\text{ind}1 = \text{MAAT}_\text{QTP})
\]

TTOP and ALT were calculated utilizing two different functions, so these two indices should be compared. For example, the two TTOP values for all stations of QTP are compared. A user can assign ind1 and ind2 to compute the ALT statistical values between the Stefan and Kudryavtsev functions. Thereafter, the statistical values are saved to the CSV file when executing the Com_Stats_QTP function. Table 3 shows all the statistical values of the selected stations.

\[
\begin{align*}
\text{Com}_\text{Stats}_\text{QTP} (\text{ind}1 & = \text{TTOP}_\text{S}_\text{QTP}, \text{ind}2 = \text{TTOP}_\text{K}_\text{QTP}) \\
\text{Com}_\text{Stats}_\text{QTP} (\text{ind}1 & = \text{ALT}_\text{S}_\text{QTP}, \text{ind}2 = \text{ALT}_\text{K}_\text{QTP})
\end{align*}
\]

A spatial trend can also be computed using the Spatial_Stat function after the regional calculation. The function simultaneously saves the spatial trend of the five indices into the NetCDF file. In addition, the function draws the animation of the spatial trend (see Section 4.4).

\[
\text{Spatial}_\text{Stat} (\text{"PIC_indices.nc"}, \text{"ALT"})
\]

### 4.4 Visualization

Station visualization can be produced by Plot_TTOP_ALT and Plot_3M. The Plot_TTOP_ALT function plots two TTOP or two ALT indices in a figure for all stations or stations with permafrost. VarName has the “TTOP” and “ALT” options, whereas SID has the “permafrost” and “all” options. The Plot_3M function draws the MAAT, MAGST, and MAGT indices. The two functions only plot these stations where permafrost exists when SID = “permafrost.”

\[
\begin{align*}
\text{Plot}_\text{TTOP}_\text{ALT} (\text{VarName} & = \text{"TTOP"}, \text{SID} = \text{"permafrost"}) \\
\text{Plot}_\text{TTOP}_\text{ALT} (\text{VarName} & = \text{"ALT"}, \text{SID} = \text{"permafrost"}) \\
\text{Plot}_3\text{M} (\text{SID} & = \text{"permafrost"})
\end{align*}
\]

The other approach of “ggplot2” was adopted to visualize the region (see Figure 4).
The indices that change over time can also be plotted through a GIF animation that uses Map_Pic (Figure 5).

The input and output of the regional calculation can be drawn using the Netcdf_Multiplot function (see Figure 6); the Netcdf_Animation function uses animation to display these values. The spatial trend can also be drawn in the Spatial_Stat apart from calculating the spatial statistics. This function draws all four indices when “VarName” has no input (see Figure 7).

5 Discussion

5.1 PIC performance

This study proposes permafrost modeling to compute the changes in the active layer and permafrost with the climate, which considers the station and region modeling over QTP. We apply the two approaches to 52 weather stations and a central region of QTP. The PIC simulation results from the 52 stations show that permafrost was detected at 12 stations based on the 52 observation stations. For QTEC, the permafrost areas began to shrink from the southern and northern parts to the central QTEC region. The permafrost, whether in permafrost stations or QTEC, continued to thaw with decreasing ALT, low surface offset and thermal offset, and high MAAT, MAGST, MAGT, and TTOP for most areas of QTP.

The PIC package computes and maps the temporal dynamics and spatial distribution of permafrost in the stations and region. The regional modeling underwent more challenges than the stations’ input data and parameters. The station calculation can estimate the long-term temporal trend of the permafrost dynamics, whereas regional calculation can estimate the temporal–spatial trend. Climate change indicates a pronounced warming and permafrost degradation in QTP (Chen et al., 2013; Cheng and Wu, 2007; Wu and Zhang, 2010; Wu et al., 2010). The simulation results show widespread permafrost degradation in QTP,
and the temporal–spatial trends of the permafrost conditions in QTP were consistent with those in previous studies (Wu and Zhang, 2010; Wu et al., 2010). In addition, the simulated TTOP and ALT that uses the Stefan and Smith functions have higher TTOP and ALT than the Kudryavtsev function. Although the overall trend of TTOP and AIT are coincidental, two different computational methods can be combined to simulate their variations. Furthermore, 16 indices can be collectively employed for a comprehensive analysis. Moreover, the station and region modeling can be integrated to evaluate the temporal–spatial evolution of permafrost in the QTP. In particular, the station modeling can be applied to validate the simulated results of the region. Moreover, regional calculation can extend from QTEC to the entire QTP.

The “for” loop is discarded, whereas the “apply” functions are used extensively to significantly lower the computation time. The current regional calculation only takes approximately 11 s. Apart from the Kudryavtsev model that requires considerable computation time (i.e., approximately 5 min), the station calculation also exhibited an improved efficiency. Therefore, PIC can be considered an efficient R package.

5.2 Advantages

Previous studies in the QTP (1) used one or two indices, such as MAAT and MAGST, to evaluate the permafrost changes (Yang et al., 2010); (2) presented a static permafrost distribution that constructed a regression analysis method through the relationship between MAGT and elevation, latitude, and slope-aspects (Nan, 2005; Yin et al., 2017); and (3) did not share the model data and codes; hence, other researchers could not validate their results and conduct further study based on previous research results (McNutt, 2014). Compared with the previous permafrost modeling in QTP, the PIC package is considerably open, easy, intuitive, and reproducible in integrating data and most of the temperature/depth-related indices. The PIC function supports the computation of multiple indices and different time periods, and the encoding mode is reusable and universal. This package can also be easily adopted to intuitively display the changes in the active layer and permafrost, as well as assess the impact of climate change. The PIC workflow is extremely simple and requires only one or two steps to obtain the simulated results and visual images. All running examples, data and code can be obtained from the GitHub repository. However, the permafrost modeling integrates satellite remote sensing data, weather and field observations, parameters, and multiple functions and models, and support dynamic change of parameters, such as vegetation and ground condition changes. Over 50
weather stations of QTP were introduced, which can approximately resolve the spatial change directions of the permafrost area. The QTEC region is an example of spatial modeling, which classifies land cover and topographic features to determine the input spatial parameters. Spatial modeling also uses the GLDAS satellite data to provide detailed information of the active layer and permafrost. The static/dynamic maps and statistical values of these indices can facilitate the understanding of the current condition of the near-surface permafrost and identify stations and ranges at considerably high risk of permafrost thawing with the changing climate and human activities. Permafrost thawing causes significant changes in the environment and characteristics of frozen-soil engineering (Larsen et al., 2008; Niu et al., 2016). A comprehensive assessment of permafrost can provide guidance regarding the future of highways and high-speed railway systems in QTP.

5.3 Limitations and uncertainties

PIC was developed with numerous indices, as well as support station and regional simulation. The PIC package can be used to estimate the frozen soil status and possible changes over QTP by calculating permafrost indices. This package has many engineering applications and can be used to assess the impact of climate change on permafrost. Moreover, this package provides observation data and provides the ability of comprehensive analysis through multiple indices. The probability of permafrost occurrence and most likely permafrost conditions are determined by computing the 16 indices. Although PIC quantitatively integrates most of these indices based on previous studies (Jafarov et al., 2012; Nelson et al., 1997; Riseborough et al., 2008; Smith and Riseborough, 2010; Wu et al., 2010; Zhang et al., 2005; Zhang et al., 2014), it still has several limitations. First, the regional calculation is one-dimensional and assumes that each grid cell is uniform without the water–heat exchange. Second, soil moisture changes at different depths affect the thermal conductivity and thermal capacity of the soil (Shanley and Chalmers, 1999; Yi et al., 2007); thus, the soil input parameters should be dynamically changed. Lastly, climate forcing has several uncertainties (Zhang et al., 2014), including input air and ground temperatures (i.e., the quality of the ground temperature in QTP is currently unreliable); thus, the regional calculation supports fewer indices than the station calculation. These deficiencies can be significant for the permafrost dynamics with environmental evolution.
6. Conclusions

An R package PIC that computes the temperature/depth-related permafrost indices with daily weather observations and climate forcing has been developed. This package is open source software and can be easily used with input data and parameters, and users can customize their own data and parameters. A total of 16 permafrost indices for stations and region are currently integrated into the R package PIC to estimate the status of the active layer and permafrost in QTP. The current functionalities also include time-series statistics, spatial statistics, and visualization. Multiple visual manners display the temporal and spatial variability on the stations and region. The package produces high-quality graphics that illustrate the status of frozen soil and may be used for subsequent publication in scientific journals and reports. The data sets of the 52 weather stations and a central region of QTP were prepared and simulated to evaluate the temporal–spatial change trends of permafrost with the climate. The simulated PIC results generally indicate that the temporal–spatial trends of permafrost conditions essentially agree with previously published studies. The PIC package has many engineering applications and can be used to assess the impact of climate change on permafrost. Additional features may be implemented in future releases of PIC to broaden its application range. In the future, the observation data of the active layer will be integrated into the PIC data set, and the output will be compared with the observation data. The PIC package can also be used to predict the future state of permafrost by utilizing projected climate forcing and scenarios. Additional functions and models will be absorbed into PIC to improve the simulation performance and perform comparative analyses with other functions and models. Parallel computation will be added to the PIC package to improve the computation efficiency. The key impact that PIC is expected to provide on the open community is an increase in consistency within and comparability among studies. Furthermore, we encourage contributions from other scientists and developers, including suggestions and assistance, to modify and improve the proposed PIC package.

Code availability

The PIC code that support the findings of this study is stored within the GitHub repository (https://github.com/iffylaw/PIC).
Data availability
The data is included in the Supplement files or GitHub repository.

Competing interests
The authors declare no competing financial interests.

Acknowledgments
This research was supported by the National Natural Science Foundation of China (41301508, 41630636). We would like to express our gratitude to the editors and anonymous reviewers for suggestions that improved this paper.

References


Yao, T., Pu, J., Lu, A., Wang, Y., and Yu, W.: Recent glacial retreat and its impact on hydrological processes on the tibetan
plateau, China, and surrounding regions, Arctic, Antarctic, and Alpine Research, 39, 642-650, 2007.


Zhang, Y., Olthof, I., Fraser, R., and Wolfe, S. A.: A new approach to mapping permafrost and change incorporating uncertainties in ground conditions and climate projections, Cryosphere, 8, 2253-2253, 2014.


### Table 1: Input data and parameters.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Meaning</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>Daily mean air temperature</td>
<td>°C</td>
</tr>
<tr>
<td>Tmax</td>
<td>Daily maximum air temperature</td>
<td>°C</td>
</tr>
<tr>
<td>Tmin</td>
<td>Daily Minimum air temperature</td>
<td>°C</td>
</tr>
<tr>
<td>GT</td>
<td>Daily mean ground temperature in 0 cm</td>
<td>°C</td>
</tr>
<tr>
<td>GT_0_MAX</td>
<td>Daily maximum ground temperature at 0 cm</td>
<td>°C</td>
</tr>
<tr>
<td>GT_0_MIN</td>
<td>Daily minimum ground temperature at 0 cm</td>
<td>°C</td>
</tr>
<tr>
<td>temp</td>
<td>Spatial daily mean air temperature</td>
<td>°C</td>
</tr>
<tr>
<td>( \lambda_t )</td>
<td>Thermal conductivity of ground in thawed state</td>
<td>W/m°C</td>
</tr>
<tr>
<td>Symbol</td>
<td>Description</td>
<td>Unit</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>( \lambda_f )</td>
<td>Thermal conductivity of ground in frozen state</td>
<td>W/m°C</td>
</tr>
<tr>
<td>( L )</td>
<td>Latent heat of fusion</td>
<td>J/m³</td>
</tr>
<tr>
<td>( \rho )</td>
<td>Dry bulk density</td>
<td>kg/m³</td>
</tr>
<tr>
<td>( W )</td>
<td>Soil water content in thawed state</td>
<td>%</td>
</tr>
<tr>
<td>( W_u )</td>
<td>Soil unfrozen water content in frozen state</td>
<td>%</td>
</tr>
<tr>
<td>( P_{sn} )</td>
<td>Period of the temperature wave, adjusted for snow melt</td>
<td>s</td>
</tr>
<tr>
<td>( C_T )</td>
<td>Volumetric heat capacity during thawing</td>
<td>J/m³</td>
</tr>
</tbody>
</table>

Figure 1: Map of the data location over QTP.
**Figure 2:** Mind map of the R package PIC.

**Table 2:** Most important user functions in the R package PIC. The equation of this table corresponds to the equation in Section 2.2.

<table>
<thead>
<tr>
<th>R function</th>
<th>Description and reference</th>
<th>Unit</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temperature-related indices</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freezing_index</td>
<td>Freezing degree-days for air and ground</td>
<td>°C</td>
<td>(4,6)</td>
</tr>
<tr>
<td>Thawing_index</td>
<td>Thawing degree-days for air and ground</td>
<td>°C</td>
<td>(3,5)</td>
</tr>
<tr>
<td>MAAT</td>
<td>Mean annual air temperature</td>
<td>°C</td>
<td>(7)</td>
</tr>
<tr>
<td>MAGST</td>
<td>Mean annual ground surface temperature (5 cm)</td>
<td>°C</td>
<td>(8)</td>
</tr>
<tr>
<td>MAGT</td>
<td>Mean annual ground temperature (at 15 m)</td>
<td>°C</td>
<td>(10)</td>
</tr>
<tr>
<td>NT</td>
<td>Thawing n-factor</td>
<td></td>
<td>(11)</td>
</tr>
<tr>
<td>NF</td>
<td>Freezing n-factor</td>
<td></td>
<td>(12)</td>
</tr>
</tbody>
</table>
Surface_Offset | The difference between the MAGST and MAAT °C
Thermal_Offset | The difference between the TTOP and MAGST °C
Vegetation_Offset | The second term (Surface_Offset) is negative, and represents the reduction in MAGST due to vegetation effects in summer (vegetation offset) °C
Nival_Offset | The first term (Surface_Offset) on the right-hand-side is positive, and represents the elevation of MAGST over MAAT due to the insulating effect of winter snow cover (nival offset) °C
TTOP_Smith | The temperature at the top of the permafrost using Smith & Riseborough function °C (13)
TTOP_Kudryavtsev | The temperature at the top of the permafrost using Kudryavtsev function °C (14)

**Depth-related indices**
Freeze_depth_Stefan | Maximum freezing depth using Stefan function m (20)
Thaw_depth_Stefan | Active layer thickness using Stefan function m (16)
ALT_Kudryavtsev | Active layer thickness (ALT) or maximum thawing depth using Kudryavtsev function m (19)

**Region**
Spatial_Pic | Spatial changes with MAAT, DDTₐ, DDFₐ and ALT m (3,4,7,16)

**Toolkit**
Com_Indices_QTP | Computing all indices for all stations of the QTP
Outlier_Process | Process the abnormal value
VLH | Computing volumetric latent heat of fusion J/m³ (2)
Convert_4_ggplot | Convert the values of TTOP & ALT to one columns
Exist_Permafrost | To determine which stations exists permafrost by TTOP values

**Statistic**
Stat | Statistic functions with more 10 methods (21,22,23)
Spatial_Stat | Spatial statistic method, just for spatial trend (24)
Com_Stats_QTP | Computing the statistic values for one or both of these indices

**Visualization**
Plot_3M | Plot MAAT, MAGST, and MAGT for all stations or a station
Plot_TTOP_ALT | Plot TTOP and ALT for all stations or a station
Figure 3: Permafrost occurrence map. Google Maps is as a base map that uses the Exist_Permafrost function. “Other” indicates the seasonal frozen soil.

Table 3: The statistical values of TTOP apply Com_Stats_QTP for the stations where permafrost exists. SD_S is the standard deviation of TTOP that uses the Stefan function, whereas SD_K is the standard deviation of TTOP that uses the Kudryavtsev function.
<table>
<thead>
<tr>
<th>Statistic</th>
<th>Tuole</th>
<th>Wudaoliang</th>
<th>Anduo</th>
<th>Maduo</th>
<th>Qingshuihe</th>
<th>Shiqu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.69</td>
<td>-0.4</td>
<td>-0.59</td>
<td>-0.9</td>
<td>-1.24</td>
<td>-1.47</td>
</tr>
<tr>
<td>Slope</td>
<td>1.11</td>
<td>1.16</td>
<td>1.2</td>
<td>1.19</td>
<td>0.93</td>
<td>0.89</td>
</tr>
<tr>
<td>R</td>
<td>0.97</td>
<td>0.96</td>
<td>0.97</td>
<td>0.97</td>
<td>0.96</td>
<td>0.86</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.94</td>
<td>0.92</td>
<td>0.93</td>
<td>0.94</td>
<td>0.92</td>
<td>0.75</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.83</td>
<td>0.86</td>
<td>0.83</td>
<td>1.24</td>
<td>1.06</td>
<td>1.5</td>
</tr>
<tr>
<td>NRMSE</td>
<td>-0.85</td>
<td>-0.34</td>
<td>-1.23</td>
<td>-0.78</td>
<td>-0.52</td>
<td>-3.17</td>
</tr>
<tr>
<td>SD_S</td>
<td>0.59</td>
<td>0.8</td>
<td>0.78</td>
<td>0.61</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>SD_K</td>
<td>0.6</td>
<td>0.66</td>
<td>0.78</td>
<td>0.66</td>
<td>0.6</td>
<td>0.69</td>
</tr>
<tr>
<td>MEF</td>
<td>-0.85</td>
<td>0.03</td>
<td>-0.06</td>
<td>-2.7</td>
<td>0.07</td>
<td>-3.09</td>
</tr>
<tr>
<td>NAE</td>
<td>0.89</td>
<td>0.39</td>
<td>1.38</td>
<td>0.86</td>
<td>0.65</td>
<td>3.35</td>
</tr>
<tr>
<td>VR</td>
<td>1.03</td>
<td>0.68</td>
<td>1</td>
<td>1.14</td>
<td>0.35</td>
<td>1</td>
</tr>
<tr>
<td>PBIAS</td>
<td>-76.13</td>
<td>-26.54</td>
<td>-108.59</td>
<td>-67.31</td>
<td>-41.42</td>
<td>-255.56</td>
</tr>
<tr>
<td>MSE</td>
<td>0.42</td>
<td>0.62</td>
<td>0.57</td>
<td>0.39</td>
<td>0.67</td>
<td>0.37</td>
</tr>
<tr>
<td>RSR</td>
<td>0.76</td>
<td>0.61</td>
<td>0.66</td>
<td>0.78</td>
<td>0.58</td>
<td>0.79</td>
</tr>
<tr>
<td>D</td>
<td>0.67</td>
<td>0.7</td>
<td>0.76</td>
<td>0.53</td>
<td>0.58</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Figure 4: TTOP that uses the Smith and Kudryavtsev functions.

Figure 5: Index changes over time for MAAT. These graphs are animated in GIF mode.
Figure 6: Regional visualization of ALT.
Figure 7: Spatial trend of MAAT, $DDT_a$, $DDF_a$, and ALT.