

王冰泉,冉有华.中国西北、西藏和周边地区1961—2020年每十年1 km季节冻土最大冻结深度数据集[J].地球科学进展,2021,36(11):1137-1145. DOI: 10.11867/j. issn. 1001-8166. 2021. 120. [WANG Bingquan, RAN Youhua. Decadal dataset of the seasonal maximum freezing depth with 1 km resolution from 1961 to 2020 in Northwest China, Tibet and surrounding area[J]. Advances in Earth Science, 2021, 36(11): 1137-1145. DOI: 10. 11867/j. issn. 1001-8166. 2021. 120. ]

# 中国西北、西藏和周边地区1961—2020年每十年 1 km 季节冻土最大冻结深度数据集\*

王冰泉<sup>1,2</sup>,冉有华<sup>1,2\*</sup>

(1. 中国科学院西北生态环境资源研究院,甘肃 兰州 730000;2. 中国科学院大学,北京 100049)

**摘要:**最大冻结深度是季节冻土热状态的重要指标,其变化对区域水循环、生态过程、气候以及工程稳定性都有重要影响。发布了我国西北五省和西藏及周边地区1961—2020年间每十年的季节冻土最大冻结深度数据集,空间分辨率为1 km。该数据集是利用支持向量机模型集合模拟产生的,以冻结指数、融化指数、降雪、降雨、太阳辐射、高程以及土壤容重为预测因子,利用基准时期(2001—2010年)气象站最大冻结深度的实测数据进行模型训练。10折交叉验证表明,模型具有良好的精度[ $(R^2 = 0.70 \pm 0.29, RMSE = (23.63 \pm 10.30) \text{ cm}, bias = (-0.77 \pm 6.01) \text{ cm})$ ]。基于实测数据的直接验证表明,1980s、1990s、2000s和2010s 4个时期模拟结果的 $R^2$ 分别为0.77、0.83、0.73和0.71, RMSE分别为27.14、22.42、21.63和23.58 cm。运用预测集合的百分位数区间评估了模拟的不确定性,结果表明模拟结果具有良好的稳定性。基于该数据集,发现1961—2020年我国西北五省和西藏地区季节冻土最大冻结深度总体呈显著下降趋势,平均每十年下降3.02 cm。1961—2010年新疆、西藏、青海、甘肃、宁夏和陕西的最大冻结深度平均每十年分别下降2.66、3.04、3.16、2.09、1.98和1.45 cm。该数据集可通过国家青藏高原科学数据中心下载(DOI: 10.11888/Geocry.tpdc.271774)。

**关键词:**第三极;高山亚洲;季节冻土;活动层;机器学习

**中图分类号:** P642.14; K903

**文献标志码:** A

**文章编号:** 1001-8166(2021)11-1137-09

## 1 引言

季节冻土指冬季冻结、夏季融化的岩层或土层<sup>[1]</sup>。我国季节冻土分布在25°N以北的地区,包括东北、华北、西北和青藏高原,约占中国陆地面积的55%<sup>[2]</sup>。土壤冻结深度是季节冻土热状态的重要指标,由于全球变暖,季节冻土的热状态极不稳定<sup>[3]</sup>,其冻结深度持续下降<sup>[4-6]</sup>,这一状态的变化改变了土

壤的热稳定性,土壤的热容量、导热率和导水率也发生变化,从而影响了陆地和大气之间的生态水文过程。土壤结构的变化也影响植被的生长状况,导致地表能量和水分的再分配<sup>[7-13]</sup>。可见,季节冻土最大冻结深度数据是理解季节冻土变化及其相关生态、水文、土壤和生物过程的基础之一。

尽管气象站有长期的季节冻土最大冻结深度

收稿日期:2021-09-27;修回日期:2021-11-01.

\* 基金项目:国家自然科学基金项目“青藏高原多年冻土退化对基础设施成本影响的统计预测研究”(编号:42071421)资助.

作者简介:王冰泉(1997-),男,河南南阳人,硕士研究生,主要从事冰冻圈遥感研究. E-mail: wangbingquan@nieer.ac.cn

\* 通信作者:冉有华(1980-),男,甘肃镇原人,副研究员,主要从事冰冻圈遥感、生态遥感与模型、遥感产品真实性检验和科学数据集成研究. E-mail: ranyh@lzb.ac.cn

的观测数据<sup>[14]</sup>,但很少有较高分辨率和长时间序列的网格数据集。而模型模拟是开展最大冻结深度制图的重要方式。目前常用的模型包括:经验模型、物理模型和统计/机器学习模型。经验模型的优点是模型结构简单、计算效率高<sup>[5,15-17]</sup>,缺点是基本假设条件难以满足<sup>[18]</sup>,导致模拟结果误差大,往往需要对模拟结果进行优化和修正<sup>[19]</sup>。物理模型是环境因素对土壤冻结深度影响过程的显式表达,它与陆面或水文模型的耦合表现出强大的模拟能力。然而,物理模型的结构比较复杂,对数据的要求高,模拟的精度和空间分辨率低<sup>[20]</sup>。

统计/机器学习模型是通过建立季节冻土最大冻结深度和环境因素之间数理统计关系实现最大冻结深度的模拟和预测。近几年,统计/机器学习模型被广泛应用于冰冻圈,例如:多年冻土制图<sup>[21]</sup>和活动层厚度模拟<sup>[22,23]</sup>等,这为季节冻土最大冻结深度的模拟提供了新的技术手段。与经验模型和物理模型不同,统计机器学习没有固定的结构,并且提供了模型误差和量化其不确定性的方法。与此同时,统计/机器学习具有一定的可移植性,已被用于预测过去和未来时期的土壤冻结深度变化。如Wang等<sup>[14]</sup>选择了空气冻结指数、土壤容重、沙土含量、黏土含量以及土壤含水量6个变量作为预测因子,结合106个钻孔观测数据训练了基准时期(2003—2010年)的支持向量机模型,并采用此模型模拟了RCP4.5情境下2040s(2041—2050年)和2090s(2091—2099年)时期的最大土壤冻结深度,表现出较好的精度。

本文在机器学习框架内采用集成模拟策略,将环境变量[冻结指数、融化指数、降雨、降雪、数字高程模型(Digital Elevation Model, DEM)、太阳辐射和土壤容重]与地面实测数据融合,模拟产生了我国西北五省、西藏和周边地区1961—2020年每十年、空间分辨率为1 km的季节冻土最大冻结深度数据集并发布共享,可为该区域季节冻土的相关研究提供数据支持。

## 2 数据生产与验证方法

### 2.1 方法

基于以前的研究,支持向量机回归模型对于季节冻土最大冻结深度的模拟具有良好的空间泛化能力<sup>[14]</sup>和可移植性<sup>[24]</sup>。由于2001—2010年季节冻土的冻结深度观测站点最多、数据最完整,因此,选择该时期作为基准时期,利用观测数据和空间预测

因子进行模型训练,并利用此模型对1961—2020年每十年的季节冻土最大冻结深度进行模拟。

首先,选择冻结指数、融化指数、降雨、降雪、太阳辐射、DEM和土壤性质(有机碳含量、土壤容重、粗粒级、粉土、黏土和沙土含量)12个地理空间数据作为潜在预测因子,其次,利用极端随机树(Extra Trees Classifier)对预测因子进行优化,选择少数重要的变量作为最后的输入数据。最后,假设数字高程模型、太阳辐射和土壤性质输入条件不变,改变不同时期的气候数据,利用2001—2010年训练的支持向量机模型模拟1961—2020年每十年的最大冻结深度。技术流程如图1所示。

#### 2.1.1 预测因子的简化

冻结指数、融化指数、降雨量、降雪量、DEM、太阳辐射和土壤性质(有机碳含量、土壤容重、粗粒级、粉土、黏土和沙土含量)都是影响季节冻土热稳定性的因素。冻结指数与季节冻土的冻结深度直接相关<sup>[17]</sup>,是影响冻土热稳定性最基本的气候指标<sup>[25]</sup>。积雪是影响冻土热状态的另一个重要因素,积雪在可见光波段内具有较高的反照率,影响地气能量平衡。土壤性质主要影响土壤的导热率和导水率<sup>[25]</sup>。另外,这些潜在的预测因子之间可能存在相关,为了解决模型可能存在的共线性问题,采用极端随机树对12个变量进行简化,并采用少数重要的变量构建支持向量机模型。基于极度随机树的特征变量选择的主要思想是选择最优的划分属性,通过计算每个特征的尼基系数使得节点处的纯度最高。*feature\_importances\_*方法返回各个特征的重要性,数值越大,该特征变量就越重要。本研究中,使用随机的特征和随机的阈值构建树,从而提供额外的随机性抑制过拟合。

#### 2.1.2 支持向量机模型集合模拟

支持向量机回归是一种非参数模型,通过确定一个误差范围( $\epsilon$ ),落在该范围内的预测值被认为是预测正确的,在给定误差间隔的大小上建立回归模型,核函数直接决定了支持向量机的性能。本研究使用最常用的核函数径向基核函数(*rbf*)<sup>[26]</sup>,*C*设置为100,*gamma*设置为0.1。该技术采用Python中的scikit-learn模块实现<sup>[27]</sup>。

#### 2.1.3 结果验证

在模型建立阶段,即2000s时期,利用该时期的实测点,基于10折交叉验证对所构建的支持向量机回归模型进行验证,采用均方根误差(Root Mean Square Error, RMSE)、决定系数(Coefficient of

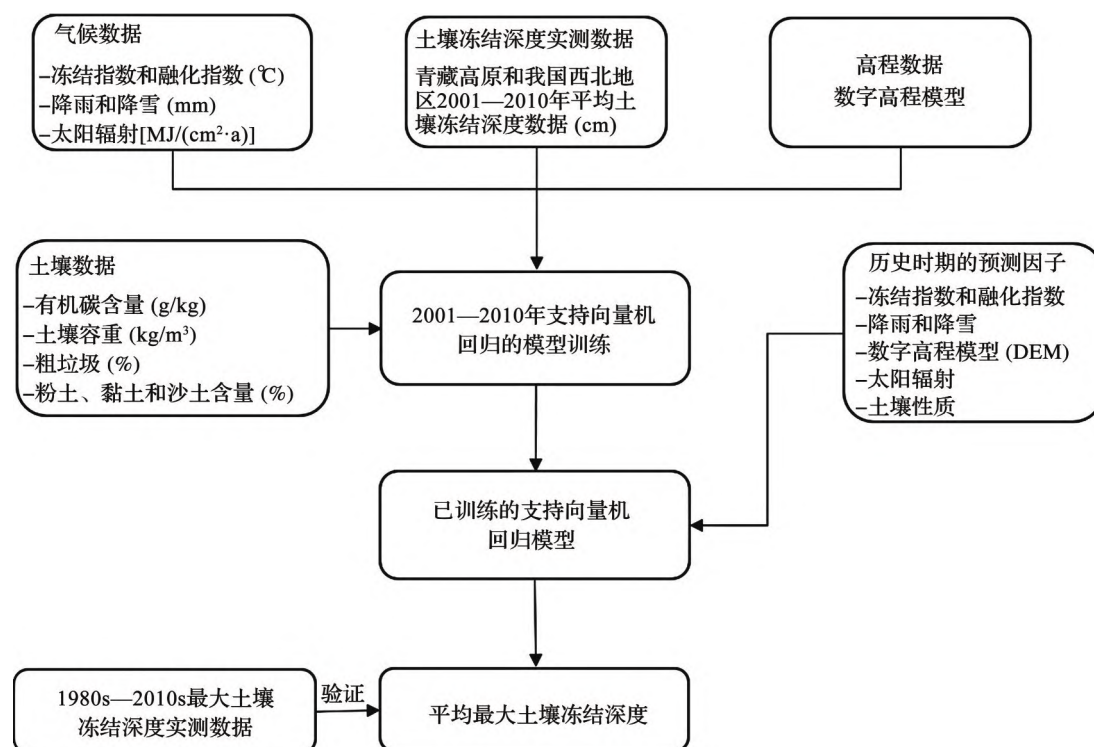


图 1 模拟季节冻土最大冻结深度技术路线图

Fig. 1 Flowchart of the process used to predict the seasonal maximum freezing depth

determination,  $R^2$ ) 和偏差 (bias) 3 个精度指标评估模型的精度。1980s、1990s、2000s 和 2010s 时期, 利用地面观测数据直接验证, 基于  $R^2$ 、RMSE 和 bias 评估模拟结果的精度。采用集合模拟百分位区间的中位数 (PI50) 和百分位区间的 95 百分位数 (PI95) 来评估各个时期模拟结果的不确定性。首先, 计算逐个像元 200 次模拟值的 95% 预测区间, 即每个像元 200 次模拟值的第 97.5 百分位数减第 2.5 百分位数 (PI)。其次, 计算整个区域 PI 值的中位数 (PI50) 和第 95 百分位数 (PI95)。PI50 对于模拟的不确定更具有代表性, 数值越小, 说明模拟的不确定越低。PI95 描述了整个区域预测值的极端值, 数值越大, 说明预测值的变异程度越大。

## 2.2 数据

### 2.2.1 季节冻土最大冻结深度实测数据

本文收集了 1981—2020 年青藏高原和我国西北地区气象站点逐日的季节冻土的冻结深度数据。首先, 查找出各站点每年的最大土壤冻结深度; 其次, 计算各站点每十年的平均最大土壤冻结深度; 最后, 1980s、1990s、2000s 和 2010s 可用的站点数量分别为 84、89、280 和 275, 2000s 时期可用数据的站点空间分布如图 2 所示。这些站点的海拔分布在 846~4 800 m, 2000s 时期平均冻结深度在 3.7~

276.6 cm, 其中, 平均土壤最大冻结深度小于 100、100~200 和大于 200 cm 的站点分别占 75.54%、22.28% 和 2.74%。基准时期 (2001—2010 年) 的实测值被用于模型的训练和验证, 在每一次模型训练中, 平均 252 个实测值用于模型训练, 剩下的实测点用于交叉验证。1980s、1990s、2000s 和 2010s 时期的实测值用于最大土壤冻结深度产品的直接验证。

### 2.2.2 气候数据

气候预测因子采用 WorldClim2.1 空间分辨率为 1 km 气候数据 (月平均温度、降水和太阳辐射) (1970—2000 年) 和空间分辨率为 5 km 的天气数据 (1960—2018 年) (<https://worldclim.org>)<sup>[28]</sup>, 已经有验证表明, 该数据集在中国区域具有较高的精度<sup>[29]</sup>。除了太阳辐射数据, 温度和降水数据不符合本研究的时间框架, 因此, 采用一种时间调整方案得到 1961—2020 年每十年 1 km 分辨率的气温和降水数据<sup>[23]</sup>。首先, 将 5 km 的天气数据使用最近邻法重采样为 1 km 以匹配气候数据的空间分辨率, 利用气候数据与重采样后的天气数据的差值的局部平滑 (3×3 的窗口) 进行调整以获得 1961—2018 年逐年的月平均温度和降水数据, 具体的调整方案可参考文献<sup>[22]</sup>。其次, 利用调整后的逐年 1 km 的气候数据计算冻结指数、融化指数、降雨和降雪。基于月



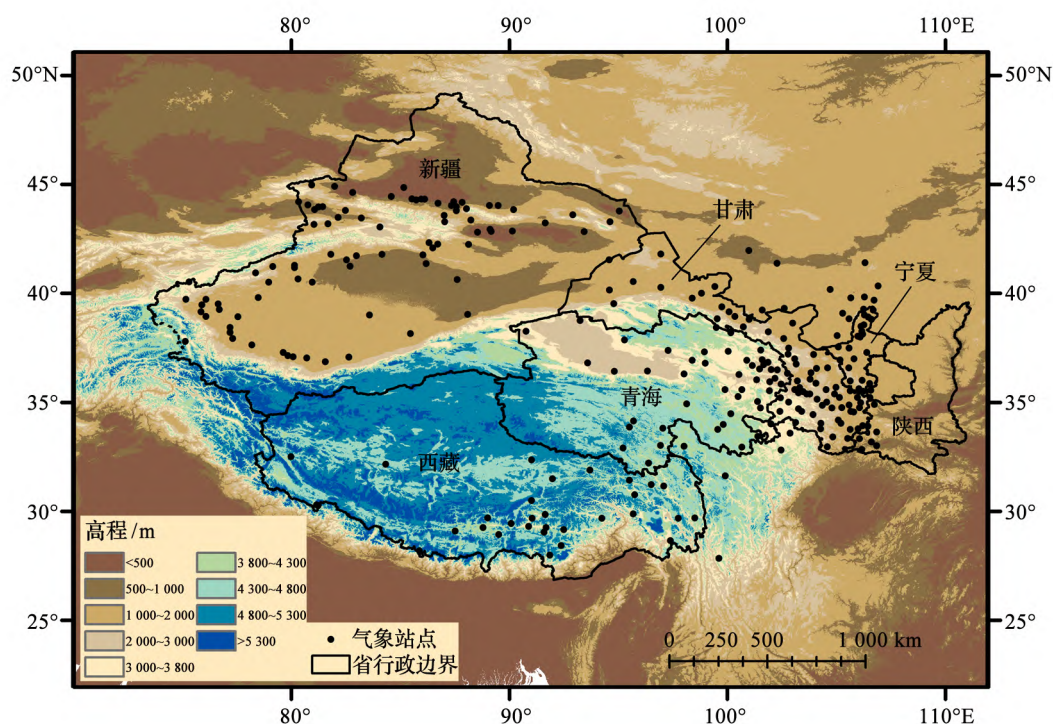


图2 最大土壤冻结深度的实测点空间分布

Fig. 2 Distribution of ground-based observation sites

平均温度计算年冻结指数和融化指数( $0^{\circ}\text{C}$ 以下和以上的月平均气温之和)<sup>[22]</sup>,基于月平均降水数据分别计算月温度大于和小于零的累积降水值代表年平均降雨和降雪。最后,计算每十年的平均冻结指数、融化指数、降水和降雪(2011—2018年的平均值代表2010s)。

### 2.2.3 数字高程模型

DEM数据来源于美国国家航天局(National Aeronautics and Space Administration, NASA)航天飞机雷达地形测绘任务(Shuttle Radar Topography Mission, SRTM)<sup>[30]</sup>。

### 2.2.4 土壤数据

土壤数据采用SoilGrids250m<sup>[31]</sup>,该数据集是由国际土壤参考资料和信息中心(International Soil Reference and Information Center, ISRIC)和世界土壤信息研究所(World Soil Information, WSI)主持研发的,在第一版本的基础之上,采用机器学习集成算法代替原来的线性回归算法<sup>[32]</sup>,空间分辨率由原来的1 km提升到250 m。该数据集综合了包括中国区域大量的地面调查数据,是目前空间分辨率最高、参数最全的土壤数据产品,已经得到了广泛应用。数据集包含了7个深度(0、5、15、30、60、100和200 cm)的土壤有机碳含量(g/kg)、容重( $\text{kg}/\text{m}^3$ )、粉

土含量、黏土含量、沙土含量(%)和粗粒级(%),本研究以7层土壤厚度作为权重,对土壤的各成分进行加权平均并重采样为1 km。

### 2.2.5 辅助数据

为了辅助最大土壤冻结深度相关统计,对冰川、水体和多年冻土区做了掩膜处理,冰川范围数据采用GLIMS(The Global Land Ice Measurement from Space)冰川数据<sup>[33]</sup>,水体范围采用GLWD(The Global Lakes and Wetlands Database)一级类别的水体数据<sup>[34]</sup>,多年冻土范围采用最近发布的高精度1 km多年冻土数据集<sup>[23]</sup>。

## 3 数据结果与验证

### 3.1 预测因子优化选择

基于极端随机树对冻结指数、太阳辐射、融化指数、降雨、降雪、DEM和土壤性质(有机碳含量、土壤容重、粗粒级、粉土、黏土和沙土含量)12个预测变量进行重要性排序,结果表明(图3),冻结指数对季节冻土最大冻结深度影响最大,土壤容重在土壤各性质中对最大冻结深度最重要,这与我们对地气热平衡的理解基本一致<sup>[4,21]</sup>。对于季节冻土的热状态的衡量,冻结指数和融化指数是基本的气候指标<sup>[25]</sup>,尤其是冻结指数,它与冻结深度呈显著正相

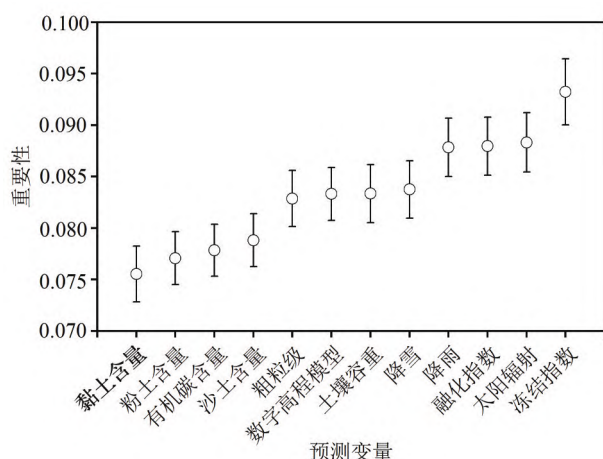


图 3 预测变量的重要性排序

Fig. 3 The importance rank of predictors

关<sup>[17]</sup>。海拔、太阳辐射和降水对冻土热状态的重要性已经成为共识<sup>[35,36]</sup>。最终,选择冻结指数、太阳辐射、融化指数、降雨、降雪、土壤容重和 DEM7 个预测变量作为模型的输入条件。

### 3.2 数据集验证结果

预测结果的验证分别从模型的精度、模拟的精度和模拟的不确定性 3 个方面进行。根据之前的研究<sup>[24]</sup>,支持向量机回归模型具有良好的精度 [ $R^2 = 0.70 \pm 0.29$ ,  $RMSE = (23.63 \pm 10.30)$  cm,  $bias = (-0.77 \pm 6.01)$  cm], 这一结论和 Wang 等的观点一致<sup>[26]</sup>。利用 1980s—2010s 每十年的实测值对模拟结果进行验证,结果表明(表 1),1980s、1990s、2000s 和 2010s 时期的  $R^2$  分别为 0.77、0.83、0.73 和 0.71,这说明基准时期的模型用于过去条件下最大冻结深度的模拟是合理的,模拟值与实测值具有较好的一致性,该模型具有良好的移植性。相较于 1991—2000 年,1981—1990 年的模拟结果误差更大, RMSE 增大了 4.72 cm。相较于 2001—2010 年,2011—2020 年的 RMSE 增加了 1.95 cm,这说明基于 2001—2010 年的支持向量机回归模型在对过去最大冻结深度空间分布预测时具有时间尺度的限制,偏离基准时期越远,模拟的精度会下降。总体而言,1960s—2010s 最大土壤冻结深度的预测结果比较稳定, PI50 和 PI95 的多年平均值(1960s—2010s)分别为 13.36 和 46.06 cm。

### 3.3 最大土壤冻结深度的时空变化分析

使用基准时期(2001—2010 年)的支持向量机回归模型,模拟过去(1960s—2010s)每十年的最大土壤冻结深度。我们假设过去时期的土壤性质、高程和太阳辐射这些条件不变,将不同时期的冻结指

表 1 1960s—2010s 中国西北和西藏地区季节冻土最大冻结深度模拟结果的精度和不确定性

Table 1 The accuracy and uncertainty of simulation results of the maximum freezing depth of seasonally frozen ground in Northwest China and Tibet from 1960s to 2010s

时期	精度			不确定性	
	$R^2$	RMSE/cm	bias /cm	PI50/cm	PI95/cm
1960s	—	—	—	13.59	47.05
1970s	—	—	—	13.56	46.81
1980s	0.77	27.14	2.55	13.56	46.67
1990s	0.83	22.42	-2.96	13.33	45.44
2000s	0.73	21.63	-0.75	12.83	44.87
2010s	0.71	23.58	1.48	13.30	45.54

注:“—”表示无数据。

数、融化指数、降雨和降雪空间数据作为模型的输入条件,1961—2020 年每十年的最大土壤冻结深度空间分布(不包括多年冻土、冰川和湖泊)如图 4 所示。

1961—2020 年我国西北五省和西藏地区季节冻土的最大冻结深度呈现显著下降趋势(图 5),平均每十年下降 3.02 cm。随着纬度和海拔的增大,最大土壤冻结深度也随之增大,新疆、西藏、青海、甘肃、宁夏和陕西平均最大土壤冻结深度分别为 99.05、109.36、140.91、93.20、72.09 和 34.16 cm。在新疆南部、青海北部和南部以及西藏北部(祁连山脉和昆仑山脉)分布着冰川,冰川边缘地区的土壤冻结深度大于 150 cm。1960s—2010s 青海的最大土壤冻结深度下降速率最快,陕西的最大土壤冻结深度下降速率最慢,平均每十年分别下降 3.16 和 1.45 cm。

## 4 数据集组成与获取

本数据集包括 1961—2020 年每十年的季节冻土最大冻结深度数据,包括 2 套数据,一套对冰川、湖泊和多年冻土进行了掩膜处理,压缩文件为“NW\_MSFD\_1”。第二套为原始数据,只对冰川和湖泊做了掩膜处理,压缩文件为“NW\_MSFD\_2”,每个文件夹中有 6 个文件,分别是 1961—2020 每十年的季节冻土最大冻结深度数据(单位:cm),例如:1960s 时期的最大冻结深度的原始数据命名为“MSFD2\_1960s.tif”。该数据集的空间范围:25°~51°N,70°~112°E,格式为 GeoTIFF,可以通过国家青藏高原科学数据中心(<http://www.tpd.cn>)下载使用,数据集名称为:中国西北、西藏及周边地区



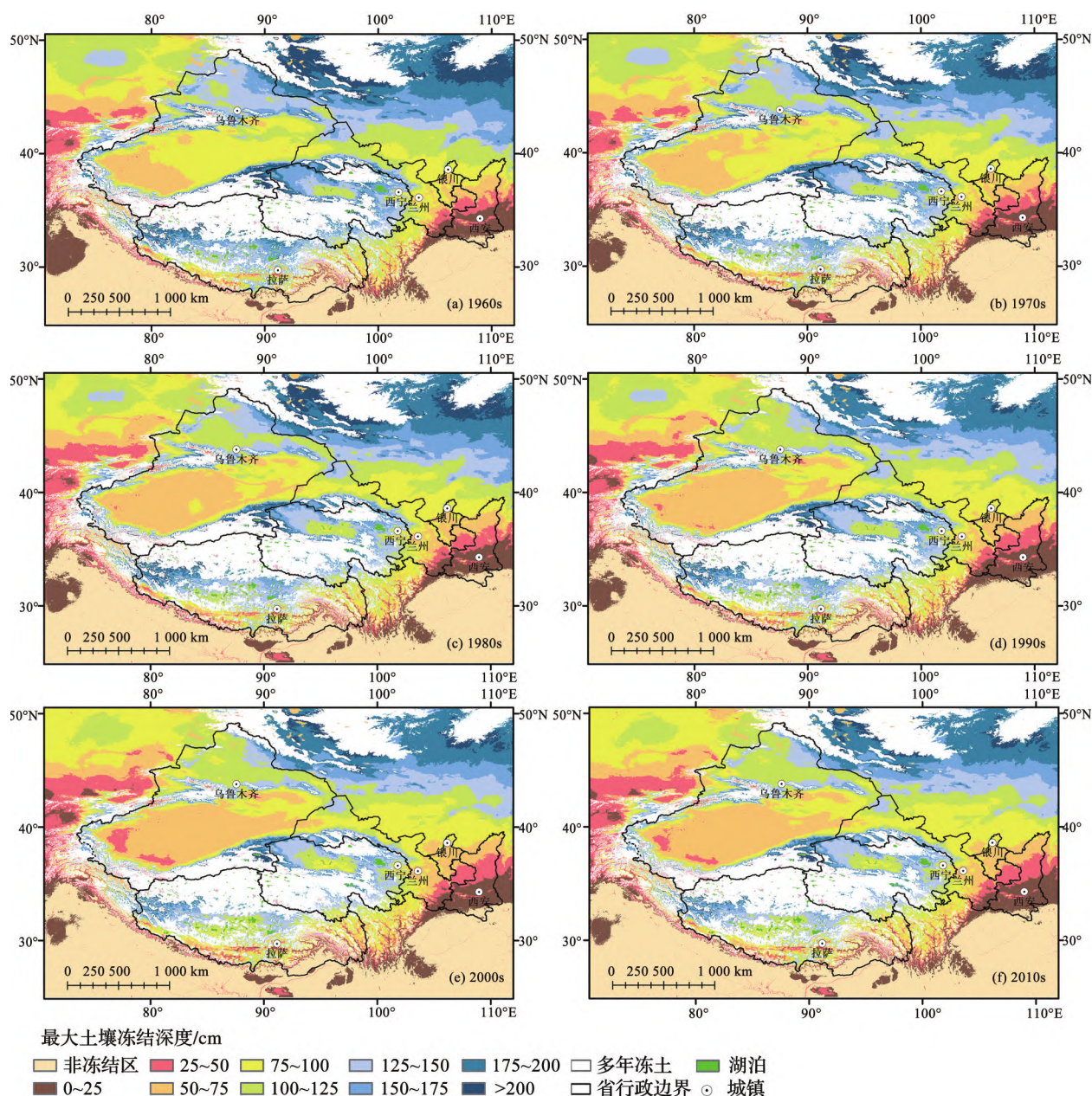


图 4 1960s—2010s 我国西北和西藏地区最大土壤冻结深度空间分布

Fig. 4 The spatial distribution of the seasonal maximum freezing depth in northwest China and Tibet from 1960s to 2010s

1961—2020 年每十年 1 km 季节冻土最大冻结深度数据(DOI: 10.11888/Geocry.tpd.271774)。

## 5 结 语

本研究在机器学习框架内利用集成模拟策略训练了基准时期的支持向量机模型,并利用此模型模拟得到了中国西北五省、西藏和周边地区 1961—2020 年每十年、空间分辨率为 1 km 的季节冻土最大冻结深度数据集。验证结果表明:2000s 时期的支持向量机回归模型具有良好的精度 [ $R^2 = 0.70 \pm 0.29$ ,

$RMSE = (23.63 \pm 10.30) \text{ cm}$ ,  $\text{bias} = (-0.77 \pm 6.01) \text{ cm}$ ], 可用于模拟过去和未来的季节冻土的最大冻结深度。最大冻结深度的模拟值和实测值具有较好的一致性,1980s、1990s、2000s 和 2010s 的  $R^2$  分别为 0.77、0.83、0.73 和 0.71。总体而言,1960s—2010s 最大冻结深度的预测结果比较稳定,PI50 和 PI95 的多年平均值(1960s—2010s)分别为 13.36 和 46.06 cm。基于该数据集,1960s—2010s 西北和西藏地区的最大土壤冻结深度呈现显著下降趋势,平均每十年下降 3.02 cm,新疆、西藏、青海、甘肃、宁夏和陕西平均

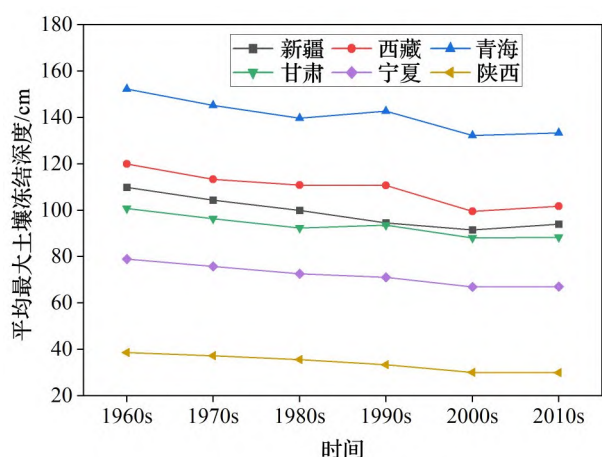


图 5 1960s—2010s 我国西北五省和西藏自治区  
平均最大土壤冻结深度

Fig. 5 The average soil maximum freezing depth in  
northwest China and Tibet from 1960s to 2010s

最大土壤冻结深度分别为 99.05、109.36、140.91、93.20、72.09 和 34.16 cm, 其中, 青海最大土壤冻结深度下降速率最快, 平均每十年下降 3.16 cm。该数据集可为中国西北、第三极、高山亚洲等相关区域的季节冻土相关研究提供数据支持。

## 参考文献 (References) :

- [1] QIN Dahe. Glossary of cryospheric science[M]. Beijing: China Meteorological Press, 2014. [秦大河. 冰冻圈科学辞典[M]. 北京: 气象出版社, 2014.]
- [2] ZHAO Lin, CHENG Guodong, DING Yongjian. Studies on frozen ground of China[J]. *Journal of Geographical Sciences*, 2004, 14(4): 411-416.
- [3] LI Xin, CHENG Guodong, JIN Huijun, et al. Cryospheric change in China[J]. *Global and Planetary Change*, 2008, 62(3/4): 210-218.
- [4] PENG Xiaoqing, FRAUENFELD O W, ZHANG Tingjun, et al. Response of seasonal soil freeze depth to climate change across China[J]. *The Cryosphere*, 2017, 11(3): 1 059-1 073.
- [5] QIN Yue, CHEN Jinsong, YANG Dawen, et al. Estimating seasonally frozen ground depth from historical climate data and site measurements using a bayesian model[J]. *Water Resources Research*, 2018, 54(7): 4 361-4 375.
- [6] LUO Siqiong, WANG Jingyuan, POMEROY J W, et al. Freeze-thaw changes of seasonally frozen ground on the Tibetan Plateau from 1960 to 2014[J]. *Journal of Climate*, 2020, 33(21): 9 427-9 446.
- [7] CHEN Boli, LUO Siqiong, LU Shihua, et al. Effects of the soil freeze-thaw process on the regional climate of the Qinghai-Tibet Plateau[J]. *Climate Research*, 2014, 59(3): 243-257.
- [8] CHRISTENSEN A F, HE H, DYCK M F, et al. In situ mea-

- surement of snowmelt infiltration under various topsoil cap thicknesses on a reclaimed site[J]. *Canadian Journal of Soil Science*, 2013, 93(4): 497-510.
- [9] JAFAROV E, SCHAEFER K. The importance of a surface organic layer in simulating permafrost thermal and carbon dynamics[J]. *The Cryosphere*, 2016, 10(1): 465-475.
- [10] LUO Siqiong, CHEN Boli, LYU Shihua, et al. An improvement of soil temperature simulations on the Tibetan Plateau[J]. *Sciences in Cold and Arid Regions*, 2018, 10(1): 80-94.
- [11] MU Cuicui, ZHANG Tingjun, WU Qingbai, et al. Carbon and nitrogen properties of permafrost over the Eboiling Mountain in the upper reach of Heihe River Basin, Northwestern China[J]. *Arctic, Antarctic, and Alpine Research*, 2015, 47(2): 203-211.
- [12] SHIKLOMANOV N I, NELSON F E. Active-layer mapping at regional scales: a 13-year spatial time series for the Kuparuk region, north-central Alaska[J]. *Permafrost and Periglacial Processes*, 2002, 13(3): 219-230.
- [13] YANAI Y, IWATA Y, HIROTA T. Optimum soil frost depth to alleviate climate change effects in cold region agriculture[J]. *Scientific Reports*, 2017, 7(1): 44860.
- [14] WANG Taihua, YANG Dawen, FANG Beijing, et al. Data-driven mapping of the spatial distribution and potential changes of frozen ground over the Tibetan Plateau[J]. *Science of the Total Environment*, 2019, 649: 515-525.
- [15] WANG Rui, DONG Zhibao, ZHOU Zhengchao. Effect of decreasing soil frozen depth on vegetation growth in the source region of the Yellow River for 1982-2015[J]. *Theoretical and Applied Climatology*, 2020, 140(3/4): 1 185-1 197.
- [16] LIU Lei, LUO Dongliang, WANG Lei, et al. Variability of soil freeze depth in association with climate change from 1901 to 2016 in the upper Brahmaputra River Basin, Tibetan Plateau[J]. *Theoretical and Applied Climatology*, 2020, 142(1/2): 19-28.
- [17] PENG Xiaoqing, ZHANG Tingjun, FRAUENFELD O W, et al. Response of seasonal soil freeze depth to climate change across China[J]. *The Cryosphere*, 2017, 11(3): 1 059-1 073.
- [18] LI Xin, KOIKE T. Frozen soil parameterization in SiB2 and its validation with GAME-Tibet observations[J]. *Cold Regions Science and Technology*, 2003, 36(1/3): 165-182.
- [19] WALVOORD M A, KURYLYK B L. Hydrologic impacts of thawing permafrost—a review[J]. *Vadose Zone Journal*, 2016, 15(6): 1-20.
- [20] RAN Youhua, LI Xin. Challenges and opportunities of permafrost mapping in China[J]. *Advances in Earth Sciences*, 2019, 34(10): 19-31. [冉有华, 李新. 中国多年冻土制图: 进展, 挑战与机遇[J]. 地球科学进展, 2019, 34(10): 19-31.]
- [21] RAN Youhua, LI Xin, CHENG Guodong, et al. Mapping the permafrost stability on the Tibetan Plateau for 2005-2015[J]. *Science China Earth Sciences*, 2021, 64(1): 62-79.
- [22] AALTO J, KARJALAINEN O, HJORT J, et al. Statistical



- forecasting of current and future circum-arctic ground temperatures and active layer thickness[J]. *Geophysical Research Letters*, 2018, 45(10): 4 889-4 898.
- [23] RAN Youhua, LI Xin, CHENG Guodong, *et al.* New high-resolution estimates of the permafrost thermal state and hydrothermal conditions over the Northern Hemisphere[J]. *Earth System Science Data Discussions*, 2021, in press. DOI: 10.5194/essd-2021-83.
- [24] WANG Bingquan, RAN Youhua. Diversity of remote sensing-based variable inputs improves the estimation of seasonal maximum freezing depth[J]. *Remote Sensing*, 2021, 13(23): 4829.
- [25] NELSON F E, OUTCALT S I. A computational method for prediction and regionalization of permafrost[J]. *Arctic and Alpine Research*, 1987, 19(3): 279.
- [26] DELUIGI N, LAMBIEL C, KANEVSKI M. Data-driven mapping of the potential mountain permafrost distribution[J]. *Science of the Total Environment*, 2017, 590: 370-380.
- [27] PEDREGOSA F, VAROQUAUX G, GRAMFORT A, *et al.* Scikit-learn: machine learning in Python[J]. *The Journal of Machine Learning Research*, 2011, 12: 2 825-2 830.
- [28] FICK S E, HIJMANS R J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas[J]. *International Journal of Climatology*, 2017, 37(12): 4 302-4 315.
- [29] PENG Shouzhong, DING Yongxia, LIU Wenzhao, *et al.* 1 km monthly temperature and precipitation dataset for China from 1901 to 2017[J]. *Earth System Science Data*, 2019, 11(4): 1 931-1 946.
- [30] JARVIS A, REUTER H I, NELSON A, *et al.* Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database [DS/OL]. 2008 [2021-08-09]. <https://srtm.csi.cgiar.org/>.
- [31] HENGL T, MENDES de Jesus J, HEUVELINK G B M, *et al.* SoilGrids250m: global gridded soil information based on machine learning[J]. *PLoS ONE*, 2017, 12(2): e0169748.
- [32] DE SOUSA L M, POGGIO L, BATJES N H, *et al.* SoilGrids 2.0: producing quality-assessed soil information for the globe [R]. *Soils and the Natural Environment*, 2020. DOI: 10.5194/soil-2020-65.
- [33] RAUP B, RACOVITEANU A, KHALSA S J S, *et al.* The GLIMS geospatial glacier database: a new tool for studying glacier change [J]. *Global and Planetary Change*, 2007, 56(1/2): 101-110.
- [34] LEHNER B, DÖLL P. Development and validation of a global database of lakes, reservoirs and wetlands[J]. *Journal of Hydrology*, 2004, 296(1/4): 1-22.
- [35] LI Xin, CHENG Guodong. The response model of high altitude permafrost to global change [J]. *Science in China Series D: Earth Sciences*, 1999, 29(2): 185-192. [李新, 程国栋. 高海拔多年冻土对全球变化的响应模型[J]. 中国科学 D 辑: 地球科学, 1999, 29(2): 185-192.]
- [36] ALA-AHO P, AUTIO A, BHATTACHARJEE J, *et al.* What conditions favor the influence of seasonally frozen ground on hydrological partitioning? A systematic review [J]. *Environmental Research Letters*, 2021, 16(4): 043008.



## Decadal Dataset of the Seasonal Maximum Freezing Depth with 1 km Resolution from 1961 to 2020 in Northwest China, Tibet and Surrounding Area\*

WANG Bingquan<sup>1,2</sup>, RAN Youhua<sup>1,2\*</sup>

(1. *Northwest Institution of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China*; 2. *University of Chinese Academy of Sciences, Beijing 100049, China*)

**Abstract:** The maximum freezing depth is an important indicator for the thermal state of seasonally frozen ground, and its changes have an important impact on the regional water cycle, ecological processes and engineering stability. This paper released a soil maximum freezing depth grid dataset for 10-year period from 1961 to 2020 in Northwest China and Tibet, with a spatial resolution of 1 km. The dataset was produced by integrating downscaled and bias corrected weather data, elevation and soil properties using a support vector machine model with 200 ensemble simulations. The 10-fold cross-validation shows that the accuracy of the support vector machine model is acceptable [ $R^2 = 0.70 \pm 0.29$ , RMSE =  $(23.63 \pm 10.30)$  cm, bias =  $(-0.77 \pm 6.01)$  cm]. Validation using in-situ data shows that the  $R^2$  for the four periods 1980s, 1990s, 2000s and 2010s are 0.77, 0.83, 0.73 and 0.71 respectively, and the RMSE are 27.14 cm, 22.42 cm, 21.63 cm and 23.58 cm respectively. The uncertainty of the simulation results is stable throughout the simulation period. Based on this dataset, we found that the soil maximum freezing depth in the Northwest China and Tibet decreased significantly between 1960s and 2020s, with an average rate of 3.02 cm per decade. The dataset can be downloaded via the National Tibetan Plateau/Third Pole Environment Data Center (DOI: 10.11888/Geocry.tpd.c.271774).

**Key words:** Third Pole; High Mountain Asia; Seasonally frozen ground; Active layer; Machine learning.

---

\* **Foundation item:** Project supported by the National Natural Science Foundation of China "Statistical prediction of the permafrost degradation impact on infrastructure future cost in Tibetan Plateau" (Grant No.42071421).

**First author:** WANG Bingquan (1997-), male, Nanyang City, Henan Province, Master student. Research areas include application of remote sensing and GIS in cryospheric research. **E-mail:** wangbingquan@nieer.ac.cn

\* **Corresponding author:** RAN Youhua (1980-), male, Zhenyuan County, Gansu Province, Associate professor. Research areas include application of remote sensing and GIS in ecohydrology and cryospheric research, science data integration, management and service.

**E-mail:** ranyh@lzb.ac.cn