

# 青藏高原冻土数据集(V1.0)

## 评估报告

冯爱霞

国家气象信息中心资料服务室

2020 年 02 月

# 目录

一、数据集基本情况 .....	1
(一) 数据源 .....	1
(二) 时空范围 .....	1
(三) 数据内容 .....	1
二、数据质量评估 .....	2
(一) 数据质量控制方案 .....	2
1、质量控制码及含义 .....	2
2、质量控制内容 .....	2
3、质量控制方法 .....	2
(二) 质量控制结果 .....	4
1、数据的缺测情况 .....	4
2、允许值范围检查结果 .....	5
3、主要变化范围检查结果 .....	6
4、内部一致性检查 .....	6
5、质量控制的总体结果 .....	7
参考文献 .....	8

## 一、数据集基本情况

### （一）数据源

目前我国已建成包含中国基本、基准气象站、一般气象站在内的 2479 个地面气象站。气象要素包括气压、气温、相对湿度、积雪、冻土、地温等。本数据集来源于《中国地面气象要素日值数据集》中的冻土要素。

### （二）时空范围

- （1）空间范围：20° N-40° N，75° E-105° E；
- （2）时间范围：1951 年 01 月-2018 年 12 月；
- （3）站点数目：20° N-40° N，75° E-105° E 空间范围内 166 个国家级地面站，其地理分布如图 1 所示。

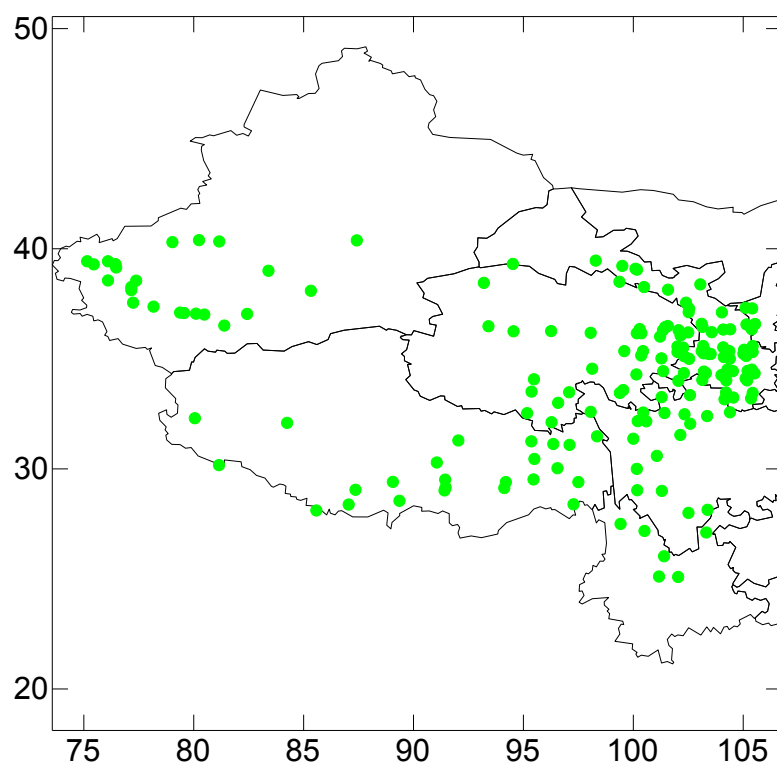


图 1 青藏高原冻土观测站空间分布

### （三）数据内容

青藏高原冻土数据资料包含 4 个要素：第一冻土层上界值和下界值，第二冻土层上界值和下界值。根据《地面气象观测规范》要求，当地面温度降到 0℃或

以下，土壤开始冻结时，每日 08 时观测一次冻土，直至次年土壤完全解冻为止。  
冻土深度不足 0.5cm 时，上、下界值均记为“0”<sup>[1]</sup>。

## 二、数据质量评估

### （一）数据质量控制方案

#### 1、质量控制码及含义

在对数据进行质量控制的过程中，随着质量控制过程的进行，需对被检验数据进行设置或者修改质量控制码，简称 QC 码，含义如下表。

表 1 数据质量控制码

质量控制码	含义
0	数据正确
1	数据可疑
2	数据错误
8	数据缺测或无观测
9	数据未进行质量控

#### 2、质量控制内容

针对第一冻土层和第二冻土层的上下界值进行质量控制的内容包括缺测检查、允许值范围检查、主要变化范围检查、内部一致性检查、人工核查与更正。

#### 3、质量控制方法

##### 3.1 缺测检查

检查第一冻土层和第二冻土层的上下界值是否为缺测数据。若为缺测数据，不再进行其他的检查，赋予该数据最终 QC 码为 8；反之赋予该数据初始 QC 码为 9，继续参与后续的质量检查。

### 3.2 允许值范围检查

检查第一冻土层和第二冻土层的上、下界值是否超过允许值范围。若超出允许值范围（允许值范围见表 2），不再进行其他的检查，赋予该数据最终 QC 码为 2；

表 2 第一冻土层和第二冻土层的上下界值的允许值范围

要素	允许上限值	允许下限值	单位
第一层冻土层上界值	450	0	cm
第一层冻土层下界值	450	0	cm
第二层冻土层上界值	400	0	cm
第二层冻土层下界值	450	0	cm

### 3.3 主要变化范围检查

计算第一冻土层和第二冻土层的上下界值各个台站每日的标准差，以及逐站逐要素逐日的距平值。若距平值介于 3 倍标准差与 5 倍标准差之间则判识为可疑，若距平值大于 5 倍标准差则判识为错误<sup>[2]</sup>。

### 3.4 内部一致性检查

通过上述步骤检查的数据参与内部一致性检查。

对于冻土同一层界值，冻土深度上界值应 $\leq$ 冻土深度下界值，否则上界值与下界值都判识为可疑；对于第二层冻土下界值应小于第一层冻土上界值，或者第二冻土层下界与第一冻土层上界均为 0 值，否则判识为两者都可疑。

### 3.5 人工核查与更正

对于上述步骤检查的疑误数据进行人工核查，核查后明确为错误的的数据判识

为错误数据且改其质控码为 2，明确为正确的数据改判其质控码为 0，无法明确质量信息的数据保留原检测结果。

## （二）质量控制结果

### 1、数据的缺测情况

第一冻土层和第二冻土层的缺测率如图 2 和图 3。从图中可知，第二冻土层缺测率同第一冻土层缺测率的情况相近。缺测率较高的站分布在高原的东南部，其他大部分站点的缺测率都在 10%以内。

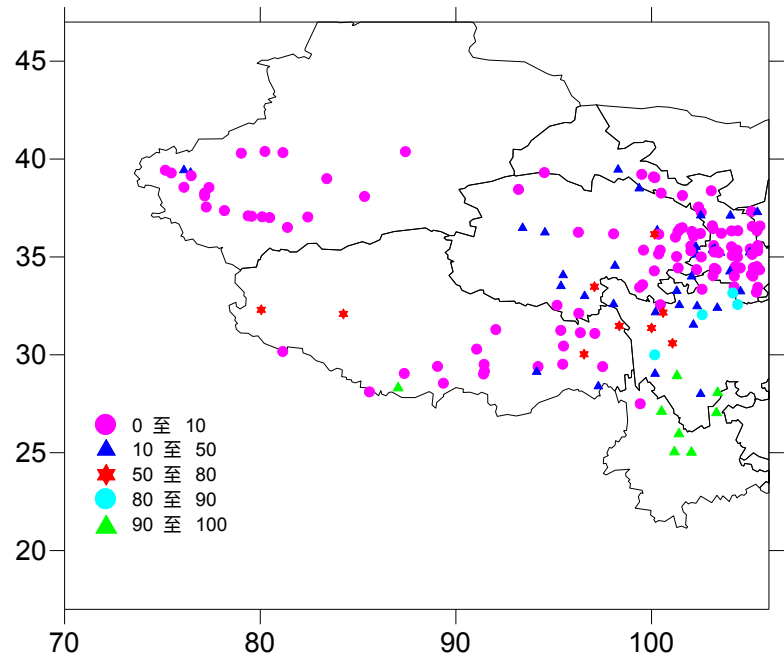


图 2 青藏高原第一冻土层缺测率地理分布图

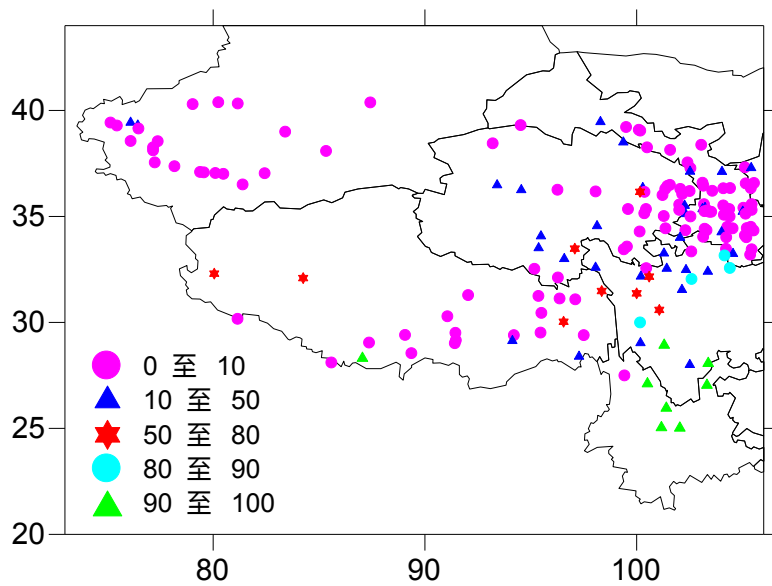


图3 青藏高原第二冻土层缺测率地理分布图

第一冻土层和第二冻土层的逐年缺测率如图4所示。从图中可知，冻土的缺测率从70年代后稳定在15%左右。

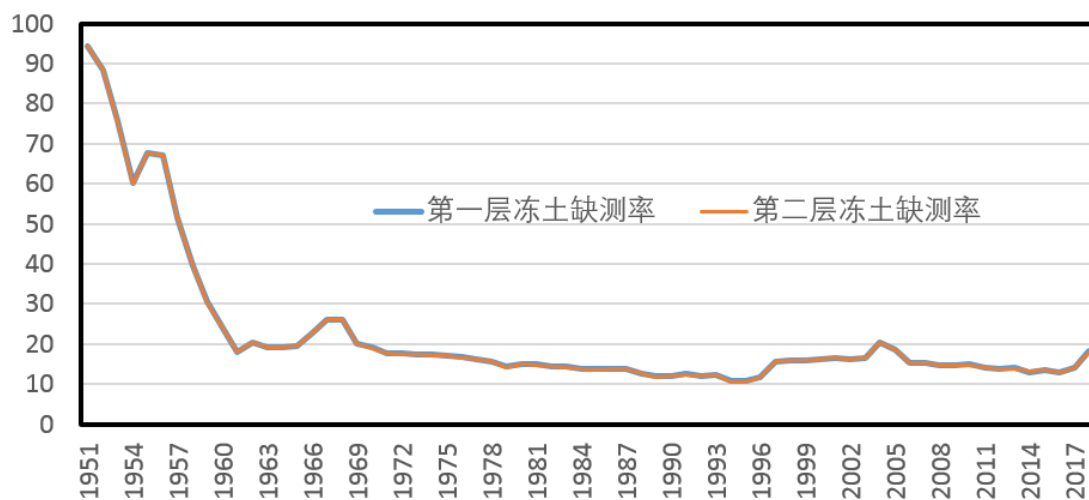


图4 1951-2018年青藏高原第一冻土层和第二冻土层数据的缺测率

## 2、允许值范围检查结果

根据第一冻土层和第二冻土层下界允许范围值 $[0, 450]$ ，第二冻土层上界范围 $[0, 400]$ ，其中第一冻土层的下界检查出共有3587条记录超越允许范围值，冻土的总记录数目为3567578，占比0.1%。经查规范和人工核查可知这些数据不是

错误数据而是实际冻土深度超出了仪器上限刻度，在上限数据基础上加 500。其他层的数据则均未超出允许范围值。

### 3、主要变化范围检查结果

各层的冻土界值的距平在大于 3 倍标准差小于 5 倍标准差以及大于 5 倍标准差的具体记录数目和占比具体见表 3。由此可知，第一冻土层的可疑率和错误率都较第二冻土层的高，下界值的可疑率和错误率较上界值的高。

表 3 各冻土层界值的允许值范围检查详情

要素	( $3\rho$ , $5\rho$ ]记录数	( $3\rho$ , $5\rho$ ]占比	大于 $5\rho$ 记录数	大于 $5\rho$ 占比
第一冻土层上界值距平	1993	0.058	902	0.026
第一冻土层下界值距平	13275	0.389	5231	0.153
第二冻土层上界值距平	47	0.001	10	0.000
第二冻土层下界值距平	538	0.016	46	0.001

### 4、内部一致性检查

对于冻土同一层界值，冻土深度上界值应 $\leq$ 冻土深度下界值，经检查冻土记录都满足此条件；对于第二层冻土下界值应小于第一层冻土上界值，或者第二冻土层下界与第一冻土层上界均为 0 值，否则判识为两者都可疑，经检查有 634 条记录不满足冻土层间的一致性，占比 0.02%。



## 5、质量控制的总体结果

经过质量控制，冻土各层界值的缺测、无观测任务、可疑以及错误情况如表 4 所示。整体上冻土数据的质量较好，数据的可疑率和错误率都较低。第一冻土层的数据完整性较好；第二冻土层无观测任务记录数较多。

表 4 冻土各层界值数据质量情况一览表

要素	无观测 任务记 录数	无观测 任务率	缺测记 录数	缺测率 (%)	可疑记 录数	可疑率 (%)	错误 记录 数	错误率 (%)
第一冻土层上界值	0	0	4521823	18.24	1993	0.056	902	0.025
第一冻土层下界值	0	0	4521823	18.24	13939	0.391	5231	0.147
第二冻土层上界值	2791086	78.23	4518983	18.16	681	0.020	10	0.000
第二冻土层下界值	2791086	78.23	4518983	18.16	538	0.015	46	0.001

## 参考文献

- [1] 中国气象局. 地面气象观测规范. 北京: 气象出版社, 2003.
- [2] 任芝花, 刘小宁, 杨文霞. 2005. 极端异常气象资料的综合性质量控制与分析, 气象学报, 63 (4): 526-533.