

多源数据融合的南极冰盖冻融数据集（1999–2019）

刘 勇^{1,2}, 周春霞^{1,2*}, 郑 雷^{1,2}, 王泽民^{1,2}

1. 武汉大学中国南极测绘研究中心, 武汉 430079;
2. 自然资源部极地测绘科学重点实验室, 武汉 430079

摘 要: 南极冰盖表面融化是全球气候变化的敏感因子, 其对于南极冰盖的物质能量平衡具有重要影响。本数据集利用微波辐射计、微波散射计和区域气候模型数据, 基于阈值法获取各个数据源的冻融时间序列, 配准后采用 Categorical Triple Collocation (CTC) 方法进行融合, 得到 1999–2019 年南极冰盖冻融产品。该数据集 4.45 km 的空间分辨率高于已发布的产品, 且结合了不同数据源的优势, 能够提供数据时段内每日南极冰盖表面融化情况。该数据集的数据类型为整型, 其中 1 代表融化, -1 代表冻结。数据集存储为.nc 格式, 压缩后的数据量为 35.7 MB。

关键词: 南极冰盖; 冻融; 数据融合; CTC

DOI: <https://doi.org/10.3974/geodp.2020.04.02>

数据可用性声明:

本文关联实体数据集已在《全球变化数据仓储电子杂志（中英文）》出版, 可获取:
<https://doi.org/10.3974/geodb.2020.05.01.V1>

1 前言

南极冰盖表面融化是全球气候变化的敏感因子, 其对于南极冰盖的物质能量平衡具有重要影响^[1]。随着全球气候变暖, 南极冰盖融化产生的径流是预估未来气候和海平面变化的很大不确定因素, 因此进一步加深对抗冰盖融化的认识意义重大。目前, 国内外发布的南极冰盖冻融产品多是采用微波辐射计数据计算得到, 其空间分辨率 (25 km) 较低, 且采用单一传感器在不同条件下 (纬度、海拔、地形等) 探测冰盖表面融化有其局限性^[2]。

本数据集基于微波辐射计、微波散射计和区域气候模型数据, 采用 Categorical Triple Collocation (CTC) 方法首先评价微波辐射计、微波散射计和区域气候模型在南极冰盖表面融化探测中的相对精度^[3], 在此基础上选取各个区域的最优观测序列, 融合生成 1999–2019 年南极冰盖冻融数据集^[4]。该数据集空间分辨率为 4.45 km, 时间分辨率为逐日, 较已经发布的产品空间分辨率高, 且采用不同的数据源, 能够将各自的优势结合起来。

收稿日期: 2020-07-21; 修订日期: 2020-12-07; 出版日期: 2020-12-24

基金项目: 国家自然科学基金 (41776200, 41941010)

*通讯作者: 周春霞 AAU-2909-2020, 武汉大学, zhoucx@whu.edu.cn

数据引用方式: [1] 刘勇, 周春霞, 郑雷等. 多源数据融合的南极冰盖冻融数据集 (1999–2019) [J]. 全球变化数据学报, 2020, 4(4): 325–331. <https://doi.org/10.3974/geodp.2020.04.02>.

[2] 刘勇, 周春霞, 郑雷等. 南极冰盖冻融数据集 (1999–2019) [J/DB/OL]. 全球变化数据仓储电子杂志, 2020. <https://doi.org/10.3974/geodb.2020.05.01.V1>.

2 数据集元数据简介

《南极冰盖冻融数据集（1999–2019）》^[4]的名称、作者、地理区域、数据年代、时间分辨率、空间分辨率、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

表 1 《南极冰盖冻融数据集（1999–2019）》元数据简表

条 目	描 述
数据集名称	南极冰盖冻融数据集（1999–2019）
数据集短名	DailyMeltingAntarctic_1999-2019
作者信息	刘勇 AAU-2576-2020, 武汉大学中国南极测绘研究中心, yonglwhu@whu.edu.cn 周春霞 AAU-2909-2020, 武汉大学中国南极测绘研究中心, zhoucx@whu.edu.cn 郑雷 AAU-3788-2020, 武汉大学中国南极测绘研究中心, zhenglei0611@hotmail.com 王泽民 AAU-3422-2020, 武汉大学中国南极测绘研究中心, zmwang@whu.edu.cn
地理区域	南极冰盖
数据年代	1999–2019
时间分辨率	逐日
空间分辨率	4.45 km
数据格式	.nc
数据量	14.13 GB（压缩后 35.7 MB）
数据集组成	南极冰盖冻融数据集（1999–2009）, 5.34 GB 南极冰盖冻融数据集（2010–2019）, 8.79 GB
基金项目	国家自然科学基金（41776200, 41941010）
出版与共享服务平台	全球变化科学研究数据出版系统 http://www.geodoi.ac.cn
地址	北京市朝阳区大屯路甲 11 号 100101, 中国科学院地理科学与资源研究所
数据共享政策	全球变化科学研究数据出版系统的“数据”包括元数据（中英文）、通过《全球变化数据仓储电子杂志（中英文）》发表的实体数据和通过《全球变化数据学报（中英文）》发表的数据论文。其共享政策如下：（1）“数据”以最便利的方式通过互联网系统免费向全社会开放，用户免费浏览、免费下载；（2）最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源；（3）增值服务用户或以任何形式散发和传播（包括通过计算机服务器）“数据”的用户需要与《全球变化数据学报（中英文）》编辑部签署书面协议，获得许可；（4）摘取“数据”中的部分记录创作新数据的作者需要遵循 10% 引用原则，即从本数据集中摘取的数据记录少于新数据集总记录量的 10%，同时需要对摘取的数据记录标注数据来源 ^[5]
数据和论文检索系统	DOI, DCI, CSCD, WDS/ISC, GEOSS, China GEOSS, Crossref

3 数据研发方法

首先评价微波辐射计、微波散射计和区域气候模型在南极冰盖表面融化探测中的相对精度及各自的优缺点。在此基础上选取各个区域的最优观测序列，生成 1999–2019 年南极冰盖逐日冻融数据集。

3.1 数据源

本数据集使用的辐射计是搭载在 DMSP（Defense Meteorological Satellite Program）卫星上的 SSM/I 传感器。该传感器最早于 1987 年发射升空，已被用来进行冰盖表面融化研究超过 30 多年^[6–7]。本数据集使用的南半球 SSM/I Level-3 EASE-Grid 亮温产品时间范围为

1999–2019 年,空间分辨率为 25 km, 时间分辨率为逐日^[8]。

本数据集使用的散射计是 C 波段的 ASCAT (Advanced Scatterometer)。在 2009 年 QSCAT 停止运行后,研究人员通常会使用 ASCAT 数据接替 QSCAT 进行积雪融化研究。本数据集使用 1999–2019 年空间分辨率为 4.45 km 的 SIR (Scatterometer Image Reconstruction) 逐日后向散射产品进行南极冰盖的表面融化研究^[9]。

RACMO2 模型以 ECMWF (European Centre for Medium-Range Weather Forecasts) 提供的 ERA-Interim 再分析资料作为驱动,能够模拟大尺度冰盖表面融化。RACMO2 已被广泛应用于南极冰盖表面物质能量平衡研究中,包括融化量、风吹雪、降水、升华等。本研究使用的 RACMO2 模型液态水含量输出产品的空间分辨率为 27.5 km, 时间范围为 1999–2019 年。

3.2 Categorical Triple Collocation (CTC) 方法原理

Triple Collocation (TC) 是一种在缺失地表观测量的情况下验证三组观测值之间精度的方法。该方法假设观测量 (M) 与真值 (X) 之间存在以下关系^[10]:

$$M_i = A_i + B_i X + \varepsilon_i \quad (1)$$

式中, A 和 B 为校准参数, ε 为随机误差, i 表示不同观测系统。TC 已被广泛应用于估算各种地表参数的误差,如土壤含水量、降水、积雪深度、表面温度、风速以及叶面积指数等。该方法假设各组观测值之间相互独立,此外,观测值与误差之间也相互独立。但是对于二值变量(如本数据集中积雪的冻融状态 1 和 -1), 式(1)要改写为:

$$M_i = X + \varepsilon_i \quad (2)$$

随机误差只有三种可能性: 0 表示观测值与真实值一致, 2 表示冻结被误判为融化, -2 表示融化被误判为冻结。显然, 对于二值变量, 观测量与误差之间有直接的关联, 因而不能使用传统的 TC 方法进行误差估算。

CTC 将传统 TC 方法进行改进, 使之能够适用于二值变量^[11]。CTC 方法所需要满足的前提条件没有传统 TC 方法那么严格。所付出的代价是, CTC 算法只能提供不同观测系统之间的相对精度, 不能估算出观测系统的绝对误差。CTC 方法通过引入平衡精度 (α) 来评价二值变量的估算精度:

$$\alpha_i = \frac{1}{2}(\psi_i + \eta_i) \quad (3)$$

式中, ψ 为正确估算为融化时的概率, 而 η 为正确估算为冻结时的概率。

静态变量的协方差矩阵 (Q) 与平衡精度 α 之间存在以下关系^[11]:

$$Q_{ij} \equiv \text{Cov}(M_i, M_j) = \begin{cases} 1 - E(E(M_i))^2, & \text{for } i = j \\ \text{Var}(X)(2\alpha_i - 1)(2\alpha_j - 1), & \text{for } i \neq j \end{cases} \quad (4)$$

这一关系推广到存在明显季节变化的非静态变量(如冻融状态)则有:

$$Q_{ij} \equiv \text{Cov}(M_i, M_j) = \begin{cases} 1 - E(E(M_i / t))^2, & \text{for } i = j \\ 4E(p(t))(1 - E(p(t)))(2\alpha_i - 1)(2\alpha_j - 1), & \text{for } i \neq j \end{cases} \quad (5)$$

式中, $p(t) \equiv pM (M = 1|t)$, t 为时间(单位为天)。CTC 算法要求各组观测系统之间的

随机误差条件独立 (即 $Pr (\varepsilon_i, \varepsilon_j | T) = Pr (\varepsilon_i | T) Pr (\varepsilon_j | T)$, $i \neq j$)。如果定义相对精度 W_i 为:

$$W_i = 2(2\alpha_i - 1)\sqrt{E(p(t))(1 - E(p(t)))}$$

(6)

式(5)中三组不同的观测系统可以形成三个等式, 则有:

$$W = \begin{bmatrix} \sqrt{\frac{Q_{12}Q_{13}}{Q_{23}}} \\ \sqrt{\frac{Q_{12}Q_{23}}{Q_{13}}} \\ \sqrt{\frac{Q_{23}Q_{13}}{Q_{12}}} \end{bmatrix}$$

(7)

由于 W_i 是以 α_i 为自变量的单调递增函数, 因此 W 的排序代表着观测系统误差由小到大的排序。将式(7)计算得到的 W 向量进行降序排列即能获取不同观测系统之间的相对误差^[3,12]。

3.3 数据集研发技术路线

本数据集研发的技术路线如图 1 所示, 首先基于阈值法获取各个数据源的冻融时间序列, 配准后再利用 CTC 方法评价相对精度并融合产品。

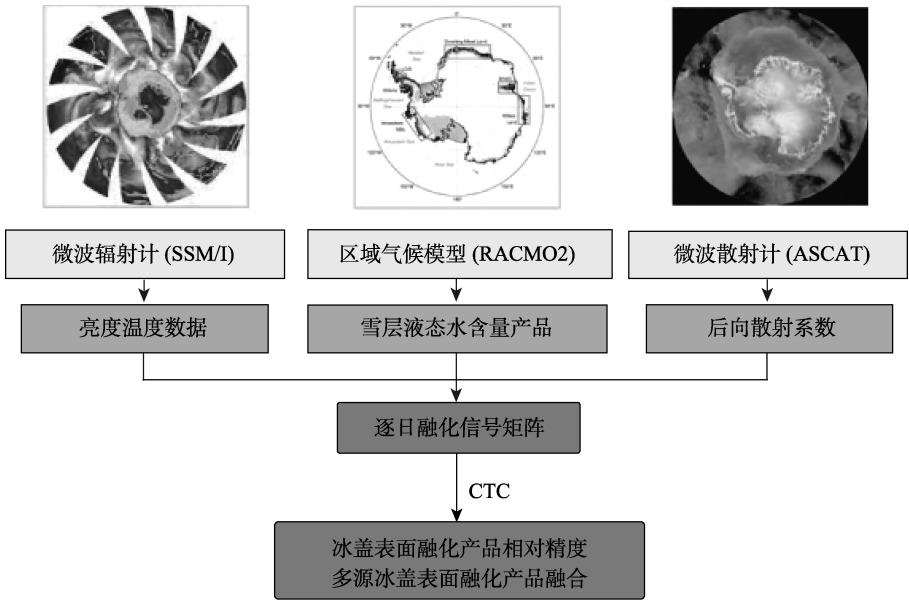


图 1 数据集研发的技术路线图

基于阈值法获取微波辐射计亮温 (T_b) 和微波散射计后向散射 (σ^0) 中的突变点。当微波亮温或后向散射超过或低于冬季平均值达到某一阈值时判断为融化信号, 其通用公式可表示为^[2]:

$$M(t) = \begin{cases} 1, Tb(t) \geq Tb_{wm} + a \\ -1, Tb(t) < Tb_{wm} + a \end{cases}, \text{微波辐射计}$$

$$= \begin{cases} 1, \sigma^0(t) \leq \sigma_{wm}^0 - b \\ -1, \sigma^0(t) > \sigma_{wm}^0 - b \end{cases}, \text{微波散射计}$$
(8)

式中, t 为时间 (d), $M(t)$ 为冻融状态 ($M = 1$ 为融化, $M = -1$ 为冻结), Tb_{wm} 为冬季 (6 月到 8 月) 平均亮温, σ_{wm}^0 为冬季平均后向散射^[3]。本数据计算时将 a 设为 30 K, 使用 SSM/I 19 GHz 水平极化亮温数据获取融化信号^[6,9,13]。SSM/I 的升降轨数据均用来进行融化探测, 当任一轨观测资料超过对应的动态阈值时就判断为融化。对于 RACMO2 模型模拟数据, 本文将液态水含量高于 0.4 mm (即 0.4 kg·m⁻²) 时判断为融化信号^[3]。

配准不同数据源的产品后, CTC 通过分解三组冰盖表面融化监测产品的 3×3 协方差矩阵获取其相对精度, 具体包括以下几个步骤: (1) 构建微波辐射计、微波散射计和 RACMO2 融化产品的协方差矩阵 Q ; (2) 基于式(7)通过协方差矩阵 Q 获取 W ; (3) 将 W 进行降序排列, 排首位的即为最优观测。将格网点的最优观测序列从对应的数据源产品中提取出来, 即可得到多源数据融合产品。

4 数据结果与验证

4.1 数据集组成

南极冰盖冻融数据集包含两个子集 “Melt1999_2009.nc” 及 “Melt2010_2019.nc”。数据集命名采用 “主题+地点+时间” 的形式命名, 子集命名采用 “主题+时间” 的形式命名。两个子集分别存放 1999–2009 年和 2010–2019 年两个时间段的南极冰盖冻融数据。

4.2 数据结果

多源数据融合的南极冰盖冻融数据集起止时间为 1999.07.01–2019.06.30, 每年 7 月 1 日至次年 6 月 30 日定义为一个融化年。数据集空间分辨率为 4.45 km, 时间分辨率为逐日, 数据集的数据类型为整型, 其中 1 代表融化, -1 代表冻结。数据集存储为 “*.nc” 格式, 压缩后的数据量为 35.7 MB。

由数据集得到的 2012–2013 年南极冰盖夏季融化随时间变化规律如图 2 所示。进入南半球夏季后, 南极冰盖开始大范围融化, 融化现象最先发生在南极半岛的北部及部分海岸地区, 随后逐步向较高纬度地区推进。12 月中旬, 南极半岛的拉森冰架和威尔金斯冰架已有较大面积融化。12 月下旬或 1 月上旬, 南极冰盖的融化达到峰值。此时大部分冰架已发生融化, 陆地冰盖也出现融化现象。随后融化面积逐步减少, 进入 2 月份后, 只有南极半岛及少部分冰架还在融化, 而后也逐渐冻结。

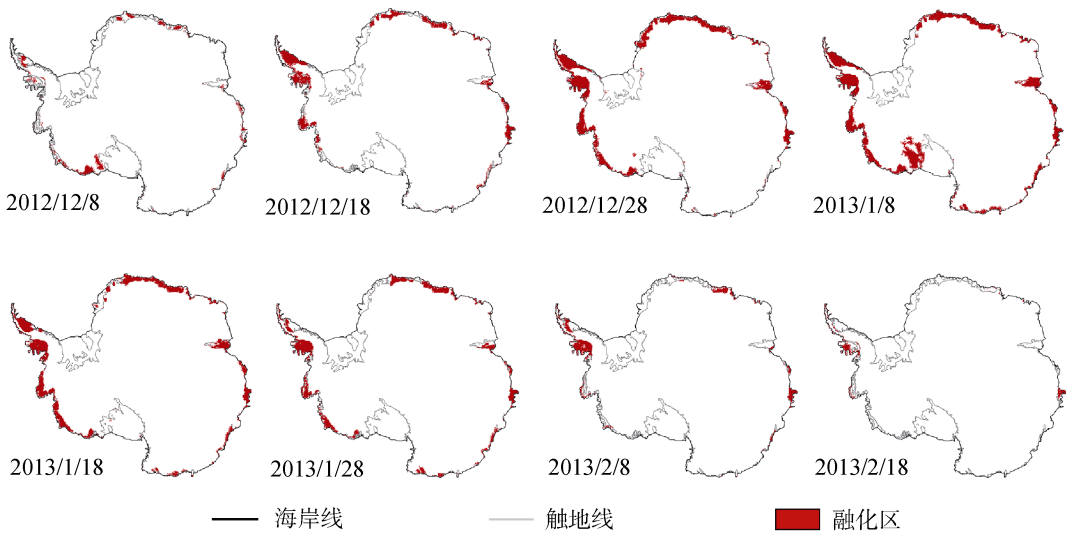


图 2 南极冰盖夏季融化随时间变化对比图

4.3 数据结果验证

由于极区雪层液态水观测资料极为稀少，而且这些资料往往并不连续。本数据集使用中国中山站（69°22'24.76"S，76°22'14.28"E）在 2016 年 7 月 1 日至 2018 年 6 月 3 日记录的近地表气温观测资料来评价积雪融化探测精度，将某日气温高于 0℃ 视为当日发生了融化，并作为真值进行评价。取中山站所对应的冻融时间序列进行验证，结果表明，产品精度可达到 92.2%。

5 讨论和总结

主被动微波传感器在南极融化探测中存在各自的优缺点，在不同区域存在互补的特征^[2]。微波散射计对雪层液态水比微波辐射计更为敏感^[7]；微波辐射计在融化剧烈的区域表现最好^[3]；模型数据由于考虑了地形因素，在卫星遥感难以工作的山地区域能够提供很好的补充^[2]。因此，本数据集结合了各个观测的优点，比单一数据源对南极冰盖冻融的探测更具优势。

综上，通过融合微波辐射计、微波散射计和区域气候模型数据，得到了高空间分辨率、高时间分辨率的南极冰盖冻融数据集。数据集经科学考察站气温数据验证，具有良好的精度。该数据集对于研究南极冰盖表面融化分布及发展过程，理解其对冰盖物质平衡的影响具有重要意义。

作者分工：周春霞、刘勇、王泽民对数据集的开发做了总体设计；刘勇采集和处理了微波及模型数据；郑雷、刘勇设计了模型和算法；刘勇做了数据验证；刘勇、周春霞撰写了数据论文。

参考文献

- [1] 王星东. 主被动微波遥感南极冰盖冻融探测[D]. 长沙: 中南大学, 2013.
- [2] 郑雷. 泛南极积雪融化微波遥感监测及多尺度驱动因子[D]. 武汉: 武汉大学, 2019.
- [3] Zheng, L., Zhou, C., Liang, Q. Variations in Antarctic Peninsula snow liquid water during 1999–2017 revealed by merging radiometer, scatterometer and model estimations [J]. *Remote Sensing of Environment*, 2019, 232: 111219.
- [4] 刘勇, 周春霞, 郑雷等. 南极冰盖冻融数据集 (1999–2019) [J/DB/OL]. 全球变化数据仓储电子杂志, 2020. <https://doi.org/10.3974/geodb.2020.05.01.V1>.
- [5] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策[OL]. <https://doi.org/doi:10.3974/dp.policy.2014.05> (2017 年更新).
- [6] Steffen, K., Abdalati, W., Stroeve, J. Climate sensitivity studies of the Greenland ice sheet using satellite AVHRR, SMMR, SSM/I and in situ data [J]. *Meteorology and Atmospheric Physics*, 1993, 51(3/4): 239–258.
- [7] Tedesco, M. Assessment and development of snowmelt retrieval algorithms over Antarctica from K-band spaceborne brightness temperature (1979–2008) [J]. *Remote Sensing of Environment*, 2009, 113(5): 979–997.
- [8] SSM/I Level-3 EASE-Grid Data Product. US National Snow and Ice Data Center-NSIDC. <https://nsidc.org/data>.
- [9] Long, D. G., Hardin, P. J., Whiting, P. T. Resolution enhancement of spaceborne scatterometer data [J]. *IEEE Transactions on Geoeence & Remote Sensing*, 1993, 31(3): 700–715.
- [10] Stoffelen, A. Toward the true near-surface wind speed: error modeling and calibration using triple collocation [J]. *Journal of geophysical research: oceans*, 1998, 103(C4): 7755–7766.
- [11] McColl, K. A., Roy, A., Derksen, C., *et al.* Triple collocation for binary and categorical variables: application to validating landscape freeze/thaw retrievals [J]. *Remote Sensing of Environment*, 2016, 176: 31–42.
- [12] Parisi, F., Strino, F., Nadler, B., *et al.* Ranking and combining multiple predictors without labeled data [J]. *Proceedings of the National Academy of Sciences*, 2014, 111(4): 1253–1258.
- [13] Jay, Z. H., Fiegles, S. Extent and duration of Antarctic surface melting [J]. *Journal of Glaciology*, 1994, 40(136): 463–475.