

基于异源数据分段特征融合的海冰厚度反演

贾飘飘¹, 韩彦岭¹, 何海洋², 王 静¹, 杨树瑚¹, 张 云¹, 洪中华¹

(1. 上海海洋大学 信息学院, 上海 201306; 2. 上海商汤智能科技有限公司, 上海 201306)

摘 要: 为了解决当前海冰厚度检测中对单一数据源依赖, 限制了海冰厚度反演精度进一步提升的问题。本文提出了一种基于异源数据分段特征融合的海冰厚度反演方法, 实验采用 Sentinel-1 合成孔径雷达数据与 ERA5 再分析数据, 通过对海冰厚度区间 $[(0, 1.5) \text{ m}, (0, 2) \text{ m}, (0, 3) \text{ m}]$ 等进行划分, 针对不同分段区间的海冰进行特征优选组合, 构建基于堆叠式集成学习的海冰厚度反演模型, 利用多个基模型和元模型之间的串并级联实现优势互补, 充分挖掘异源特征与海冰厚度之间的隐藏关联, 实现分段海冰厚度的精确反演。结果显示, 相比其他传统机器学习方法, 该方法在不同分段区间获得了较好的整体反演效果, 其中, $(0, 1.5) \text{ m}$ 区间表现最佳, 决定系数 (R^2) 达 0.923, 均方根误差 (RMSE) 低至 0.089 m。研究表明, 通过分段特征优选与异源数据融合可有效提高海冰厚度反演精度, 验证了本文提出的堆叠式集成学习模型在异源数据融合中的优势。本研究可为实现海冰厚度高精度反演提供新方法。

关键词: 海冰厚度; 异源数据融合; 分段特征融合; 堆叠式集成学习; 北极

中图分类号: P 715; P 731.15 **文献标志码:** A

海冰具有较高的反射率, 这使得地表吸收的能量显著减少, 进而对大气和海洋之间的能量分配产生重要影响^[1]。在海冰诸多特性中, 海冰厚度这一关键物理参数的动态变化, 在上述能量分配以及物质和能量交换过程中扮演着极为重要的角色, 不仅影响大气和海洋之间的物质交换和能量平衡^[2], 而且对各地的气温、降水等气候要素产生直接或间接作用, 为全球气候分析提供不可或缺的数据支持, 因此对海冰厚度的监测和研究十分重要^[3-4]。

早期的海冰厚度监测大多采取现场测量, 使用电钻在冰面上打孔, 然后进行实地测量。这种方法虽然能够获取准确的海冰厚度, 但是需要耗费巨大的人力和物力, 且其测量范围小, 难以进行大面积测量^[5]。随着科技的发展, 出现越来越多的非直接接触的测量方法, 例如: 仰视声呐测量、电磁感应、船测摄影测量和卫星遥感测量。卫星遥感朝着大范围、高精度方向不断发展, 给海冰厚度反演提供了技术手段^[6-7], 其中, 基于合

成孔径雷达 (Synthetic aperture radar, SAR) 的主动微波遥感由于具有高分辨率、全天时和全天候的特点, 且不受雨雾等天气因素的影响, 成为海冰厚度检测的理想数据来源^[8-9]。

目前遥感海冰厚度反演的研究主要包括以下几类: (1) 积分方程法 (Integral equation method, IEM), WAKABAYASHI 等^[10]采用积分方程法反演海冰厚度, 研究发现, 使用 SAR L 波段中的后向散射系数 VV (Vertical transmit - Vertical receive) 和 HH (Horizontal transmit - Horizontal receive) 的比值, 不仅对海冰表面的介电常数极为敏感, 而且与海冰厚度呈现出较高的相关性^[10]; KIM 等^[11]使用 RADARSAT-2 (C 波段) 和 TerraSAR-X (X 波段) 的 SAR 数据采用积分方程法和扩展 Bragg 散射模型反演海冰厚度, 研究发现, 冬季海冰厚度与共极化和交叉极化比之间存在很强的相关性, 这表明共极化和交叉极化比是估算冬季海冰厚度的有效特征。(2) 海冰电磁 EM (Electromagnetic) 散射模型与海冰生长模型

收稿日期: 2025-01-13 修回日期: 2025-04-25

基金项目: 国家自然科学基金 (42176175, 42271335); 国家重点研发计划 (2019YFD0900805)

作者简介: 贾飘飘 (1995—), 女, 硕士研究生, 研究方向为遥感海冰厚度。E-mail: m220951613@st.shou.edu.cn

通信作者: 韩彦岭, E-mail: ylh@shou.edu.cn

版权所有 ©《上海海洋大学学报》编辑部 (CC BY-NC-ND 4.0)

Copyright © Editorial Office of Journal of Shanghai Ocean University (CC BY-NC-ND 4.0)

<http://www.shhydx.com>

相结合,ZHANG 等^[12]采用了海冰 EM 散射模型与海冰生长模型相结合的方法,评估了使用 SAR 在不同波段(C 波段,X 波段,L 波段)、极化方式和入射角等多个方面对渤海薄海冰的反演效果,其结果表明不同波段、极化方式和入射角对海冰厚度的影响各有不同。(3)线性回归方法,其基本思想是建立特征值和海冰厚度之间的回归方程,刘眉洁等^[13]的研究首次聚焦于单特征入射角在海冰厚度反演中的应用,使用线性回归方法证明了使用单特征入射角反演海冰厚度比传统利用后向散射系数的误差要小,其相关系数为 0.81,平均绝对误差(Mean square error, MAE)为 0.31;刘立剑等^[14]创新性地采用多元线性回归结合极化分解技术,对基于 C 波段 PolSAR 遥感图像进行极化分解,得到多组特征,证明了多特征反演海冰厚度的优越性;于淼等^[15]则另辟蹊径,将纹理特征引入海冰厚度反演研究,使用多元线性回归方程,从 RADARSAT-2 SAR 影像中,使用灰度共生矩阵计算纹理特征,利用纹理特征与海冰厚度之间的关系,选取相关性较高的纹理特征,进行纹理特征融合反演海冰厚度,进一步提高了海冰厚度的反演精度。(4)利用新的机器学习方法,如神经网络、随机森林、增加注意力机制的卷积神经网络等方法,HERBERT 等^[16]采用神经网络方法挖掘 L 波段亮温数据与海冰厚度隐藏的联系,分别对厚度小于 0.6 m 的薄冰和厚度小于 2.5 m 厚冰建立模型,获得了较好的效果;SHAMSHIRI 等^[17]采用随机森林的方法,从 Sentinel-1 中提取后向散射系数(HH, HV)和月份进行回归反演,其反演精度决定系数(R^2)为 0.72, MAE 为 0.15;LIANG 等^[18]运用具有注意力机制的卷积神经网络,借助热力学特征反演每日海冰厚度,证明了热力学特征反演海冰厚度的可行性。上述研究方法中,前 3 种方法主要利用海冰遥感特征与海冰厚度之间的线性拟合进行海冰厚度反演,针对物理性质复杂的厚冰其反演精度受到一定限制,随机森林、集成学习等新的机器学习方法凭借在解决大数据量、非线性问题方面的优势,为进一步提高海冰厚度反演精度提供了技术手段。然而,目前这些方法在解决实际问题时仍存在各自的局限性,如何充分整合各类方法的优势,进一步提高海冰厚度反演精度,仍是该领域需要考虑的重要问题。

通过对以上研究分析发现,目前海冰厚度反演大多数利用单一数据源对海冰厚度进行整体范围的研究,单源数据的局限性以及不同厚度海冰形成过程、环境影响因素和物理特性的显著差异在一定程度上限制了海冰厚度反演精度的进一步提升。基于此,本文提出基于堆叠式集成学习的异源数据分段特征融合海冰厚度反演方法(Stacking ensemble learning sea ice thickness inversion method based on heterogeneous data segmented feature fusion, Stacking_ELSITM),利用异源数据丰富的特征信息(如 SAR 数据的极化、纹理特征,以及 ERA5 数据中的温度、辐射、反照率等气象特征),充分挖掘有利于不同分段海冰厚度反演的主动微波遥感数据特征和热力学特征,构建针对不同厚度分段区间的优化特征组合,实现异源数据的优势互补,然后利用随机森林、XGBoost、极端随机树和深度森林在处理大数据量、非线性问题方面的优势,基于 Bootstrap 采样策略构建堆叠式集成学习模型,通过多个基模型深度挖掘多源特征与海冰厚度之间的隐藏关联,最后将多个基模型的输出结果输入元模型获得海冰厚度反演结果,本文方法通过异源数据融合和分段特征优选等方式进一步提高了不同分段区间的海冰厚度反演精度,获得了较好的整体反演效果。

1 材料与方法

1.1 研究区域

本文的研究区域位于北冰洋边缘的波弗特海,北接北冰洋中心海域,南靠美国阿拉斯加,东邻班克斯岛,西接楚科奇海,65°N~80°N 和 120°W~165°W 海域。波弗特海属于极地地区,受极地气候的影响,常年气候寒冷,海面几乎全年结冰,只有在 8—9 月,才会出现很少的海面无冰的状态。波弗特海水由陆地的河水、大西洋和太平洋的海水汇集而成,由于河水和海水在温度和盐度方面存在差异,从而汇集在波弗特海的不同深度,致使温度在竖直方向出现变化,对海冰厚度变化产生影响。

1.2 数据与预处理

1.2.1 Sentinel-1 遥感卫星数据

Sentinel-1 是由欧洲空间局发射的卫星,其目的是用于监测陆地和海洋。本文使用的是

Sentinel-1 卫星合成孔径雷达(SAR)C 波段超宽幅(Extra wide,EW)模式的 GRD 产品数据,具有中等空间分辨率 90 m,像素分辨率 40 m×40 m,双极化模式(HH 和 HV)^[19]。Sentinel-1 由 2 颗卫星组成,分别为 Sentinel-1A 和 Sentinel-1B。Sentinel-1A 于 2014 年 4 月 3 日发射升空, Sentinel-1B 于 2016 年 4 月 25 日发射成功,这 2 颗卫星位于同一轨道,轨道相位差为 180°,可以实现 6 d 完成 1 次对全球的完整成像,对北极地区的重访周期不到 1 d。在本文的研究中,我们共选用了 5 景 Sentinel-1 图像,具体包括:2017 年 4 月 2 日 Sentinel-1A 获取的 2 景、2017 年 4 月 6 日 Sentinel-1B 获取的 2 景以及 2017 年 4 月 8 日 Sentinel-1A 获取的 1 景。图 1 为 5 景 SAR 图像在波弗特海区域的位置示意图,图 2 给出了 5 景 SAR 图像的原始彩色图(图 2a~2e)及其相应的

HH(图 2f~2j)和 HV(图 2k~2o)灰度图。

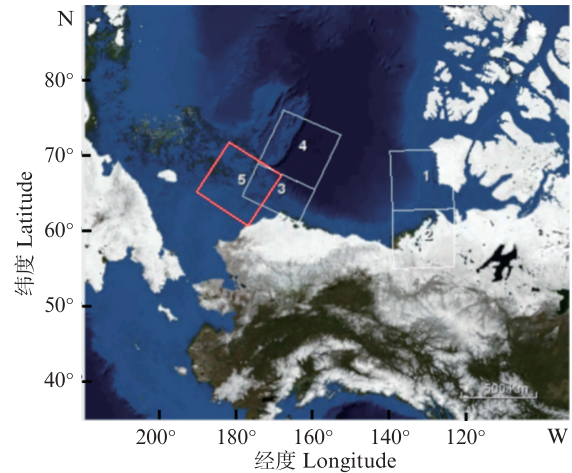


图 1 SAR 图像在波弗特海覆盖区域示意图
Fig. 1 Schematic representation of the SAR images in the area covered by the Beaufort Sea

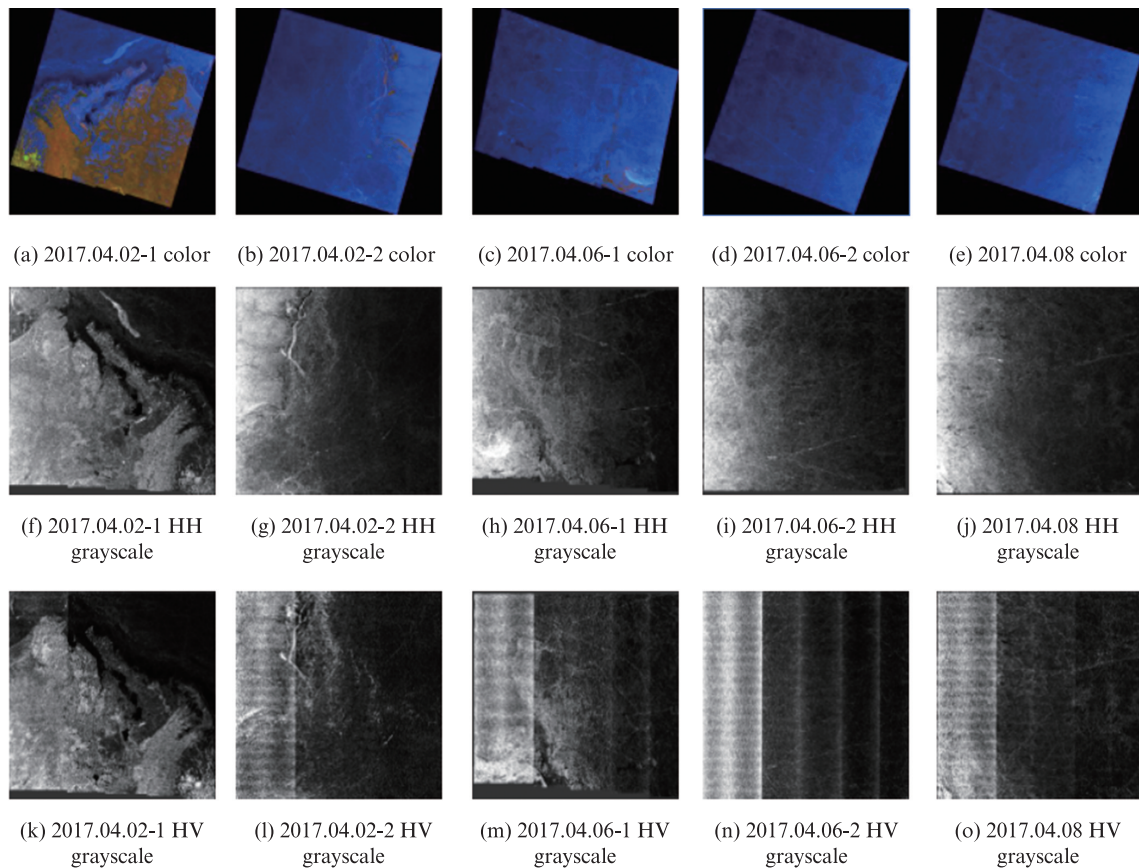


图 2 5 幅合成孔径雷达图像的原始彩色图及其相应的 HH 和 HV 灰度图
Fig. 2 Original colour map of 5 SAR images and their corresponding HH and HV grayscale maps

1.2.2 ERA5热力学数据

欧洲中期天气预报中心(European centre for medium - range weather forecasts, ECMWF)推出的全球再分析数据集 ERA5(The fifth generation of ecmwf reanalysis) (<https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels%3Ftab=form?tab=overview>)提供了从1979年到现在的一整套高质量数据集,空间分辨率为0.25°,时间分辨率是小时级别。其特征变量共240种,包括气温、气压、风、辐射、云量、降雨、湿度等,为全球气候的研究和各学科的应用提供宝贵的资料^[20]。

海冰厚度与热力学因素之间存在着紧密且复杂的关系。从热力学角度来看,多种因素共同作用影响着海冰厚度的变化,在海冰的生长与消融过程中,太阳辐射起着关键作用,当海冰吸收

的太阳辐射量超过其释放的能量时,海冰厚度就会随之减少,新形成的海冰甚至可能因此融化。感热、潜热辐射在海冰吸收热量以后导致海冰表面温度增加,因此感热、潜热辐射与气温和海表温度有一定的关系;风通过推动海冰运动及改变热量交换条件影响潜热辐射的变化;雪层温度能够影响海冰的融化从而导致海冰厚度变化;另外,海冰厚度和反射率之间存在一定关系,研究表明,开阔水和海冰的反射率有很大的差异,随着海冰厚度的变化,海冰的反射率为0.5~0.9^[21],这些热力学因素相互关联,共同影响着海冰厚度的变化。

综上,影响海冰厚度的主要热力学因素包括感热辐射、潜热辐射、2 m气温、海表温度、10 m高度的纬向风(u)、10 m高度的经向风(v)、雪层温度、预测反射率等,具体如表1所示。

表1 热力学特征参数
Tab. 1 Thermodynamic characteristic parameters

变量名 Variable name	英文简写 English abbreviation	单位 Unit
感热辐射 Surface sensible heat flux	SLHF	W/m ²
潜热辐射 Surface latent heat flux	SSHF	W/m ²
2 m 气温 2 m temperature	T2M	K
海表温度 Skin temperature	SKT	K
10 m 高度的纬向风(u) 10 m u-component of wind	U10	m/s
10 m 高度的经向风(v) 10 m v-component of wind	V10	m/s
雪层温度 Temperature of the snow layer	TSN	K
预测反射率 Forecast albedo	FAL	%

1.2.3 海冰厚度验证数据

本研究使用的海冰厚度数据来自于阿尔弗雷德韦格纳研究所(Alfred wegenerinstitute, helmholtzcentre for polar and marine research, AWI) 在2017年3—4月的一个北极气候模拟项目,其数据经过50次的调查,利用机载电磁感应器(Electronic warfare, EW)进行实际测量并计算获得,其数据包括经纬度、测量的日期、时间和海冰厚度等信息^[22],下载链接:(<https://doi.pangaea.de/10.1594/PANGAEA.924848>)。

1.2.4 数据预处理

对Sentinel-1 图像进行的预处理包括轨道校正、热噪声去除、辐射校准、多普勒地形校正和分贝化,其中,轨道校正是用精确轨道文件更新卫星轨道信息;热噪声去除是归一化反向散射信号,剔除影像中的热

噪声;辐射校准主要是修正数据获取及传输中的系统性、随机性失真;多普勒地形校正是消除多普勒效应的影响;分贝化是对后向散射系数作对数变换,便于数据分析。具体预处理流程如图3所示。

纹理特征是指图像中同质现象的视觉特征,能够表征不同类别地物的纹理、颗粒等细节特征,海冰在不同的生长阶段,不同厚度的海冰其表面纹理特征会呈现出显著的差异,通过提取纹理特征并分析和挖掘其与海冰厚度之间的潜在关系,有助于提高海冰厚度反演的准确性^[23]。灰度共生矩阵是研究图像纹理的空间分布和灰度关系的有效方法,它通过计算灰度图像中各个像素灰度值之间的空间关系,利用生成的共生矩阵特征值来描述图像中的纹理特征,具体包括图像灰度在方向、相邻间隔、变化幅度上的信息^[24]。

本文基于灰度共生矩阵提取纹理特征,由于HH极化对地物表面粗糙程度的敏感度高于HV极化^[15],因此采用HH极化方式进行纹理特征提取,纹理特征提取中的4个重要参数是窗口大小、

方向、位移量d和灰度量化级别,本文的参数分别采用9×9的窗口、全方向入射角、位移是4和32的灰度级^[25],根据Haralick定义的矩阵进行纹理特征值计算^[15],详见表2。

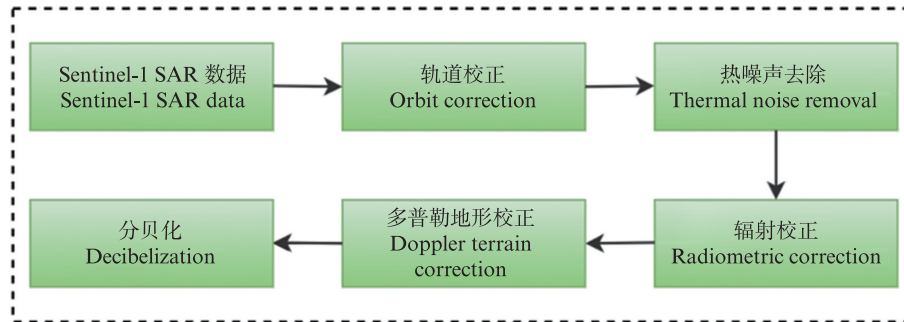


图3 Sentinel-1数据预处理流程图

Fig. 3 Flow chart of data pre-processing of Sentinel-1

表2 GLCM 纹理特征值表

Tab. 2 GLCM texture feature values table

纹理特征名称 Texture feature name	公式 Formula	纹理特征名称 Texture feature name	公式 Formula
对比度 Contrast ratio	$Con = \sum_{i,j=1}^{Ng} 1(i-j)^2 P_{ij}$	最大概率 Maximum probability	$MAX = \max P_{ij} $
相异性 Diversity	$Dis = \sum_{i,j=1}^{Ng} P_{ij} i-j $	熵 Entropy	$Ent = -\sum_{i,j=1}^{Ng} P_{ij} \cdot \ln(P_{ij})$
同质性 Homogeny	$Hom = \sum_{i,j=1}^{Ng} \frac{P_{ij}}{1+(i-j)^2}$	均值 Mean value	$Mea = \frac{1}{Ng^2} \sum_{i,j=1}^{Ng} P_{ij}$
二阶矩 Second moment	$ASM = \sum_{i,j=1}^{Ng} (P_{ij})^2$	方差 Variance	$Var = \sum_{i,j=1}^{Ng} P_{ij} (i-u)^2$
能量 Energy	$Ene = \sqrt{\sum_{i,j=1}^{Ng} (P_{ij})^2}$	相关性 Variance	$Cor = \frac{\sum_{ij}(ij)P_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y}$

注: μ 和 σ 是 P 的均值和方差。

Notes: μ and σ are the mean and variance of P .

本研究首先根据飞机走航数据的时间,选取与走航数据相同时间的 Sentinel-1 卫星数据和 ERA5 数据进行时间匹配;由于 Sentinel-1 数据的像素分辨率为 40 m,而 ERA5 数据的空间分辨率为 0.25°,因此采用双线性插值法将 ERA5 数据重采样插值至 40 m,与 Sentinel-1 数据的像素分辨率一致;然后根据 Sentinel-1 卫星影像的经纬度,选取位置最接近的飞机走航数据进行经纬度匹配,并将该位置的走航数据作为海冰厚度的实测值,从而形成遥感数据集。同样,ERA5 数据的数据匹配过程与 Sentinel-1 数据的数据匹配过程相同,重复上述过程,形成热力学数据集。本文提取了 2017 年 4 月 2 日、4 月 6 日以及 4 月 8 日的异源数据,共得到 163 796 条样本数据,样本数据集的厚度分布如图 4 所示。从图上看,海冰厚度分

布范围很大,其中厚度值低于 0 m 的被认为是异常值并进行剔除,剔除异常值后数据集共包含 161 472 条样本数据,厚度分布如图 5 所示。

1.3 基于堆叠式集成学习的海冰厚度反演总体框架

本文建议方法的主要思想是充分挖掘异源数据特征,分析不同厚度分段区间的海冰特性,进而构建针对不同分段区间的海冰厚度反演优化特征组合,利用随机森林、XGBoost、极端随机树和深度森林的优势构建基模型,深入挖掘不同分段区间的海冰特征和厚度之间的复杂关联,然后构建元模型并整合多个基模型的训练结果进行再次训练,通过堆叠式集成学习的方式提高模型的泛化能力。

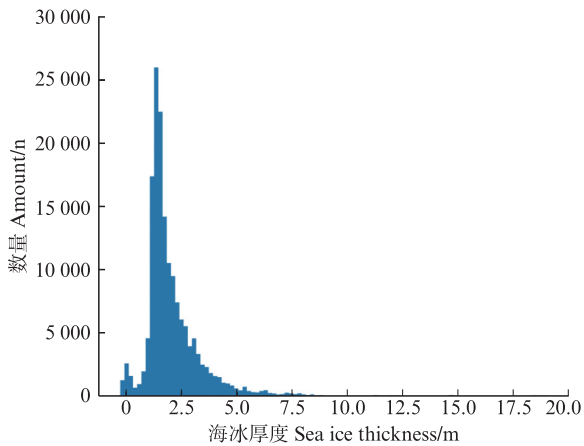


图4 海冰厚度实际分布图

Fig. 4 Actual distribution map of sea ice thickness

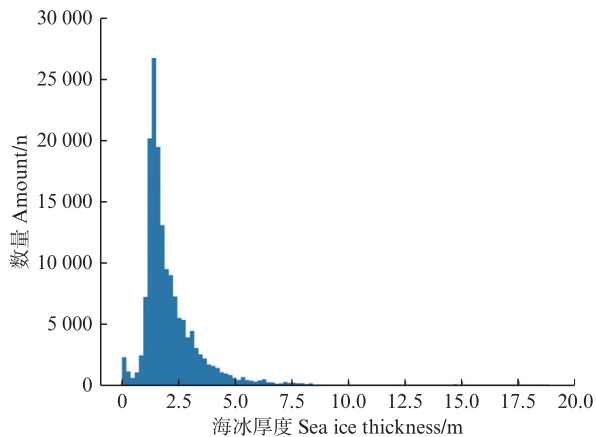


图5 剔除异常数据后的海冰厚度分布图

Fig. 5 Sea ice thickness distribution plot after removing outlier data

基于堆叠式集成学习的海冰厚度反演总体框架如图6所示,主要包括三个部分:第一部分是数据预处理,首先对SAR数据进行预处理,提取入射角、后向散射系数、纹理等特征,对ERA5数据进行数据整合和处理,提取热力学特征,然后根据时间和经纬度与飞机走航数据(作为验证数据)进行时空匹配,从而建立数据集;第二部分是方法主体部分,针对海冰的不同厚度分段区间,对遥感特征和热力学特征进行重要性分析,并在分段区间内对这两类异源数据进行特征优选组合,通过构建基于堆叠式集成学习的反演模型进行深度特征提取和特征融合,实现海冰厚度反演,并与其他经典机器学习方法进行对比分析;第三部分是精度评估,利用回归系数、MSE、RMSE等对海冰厚度反演结果进行精度评估。

1.3.1 堆叠式集成学习海冰厚度反演模型

Stacking_ELSITM模型由3个基学习器组成的基模型和1个元模型构成,融合了Bagging和Boosting集成学习的思想,通过多个学习器的串并联组合构造堆叠式集成学习模型,利用3个不同基学习器的特性,充分挖掘不同分段区间的海冰特征和厚度之间的隐藏关联,实现海冰特征的深度融合和基学习器的优势互补,从而提升海冰厚度反演的整体性能^[26-27]。具体模型结构如下图7所示,第一层以随机森林、XGBoost和极端随机树构建基模型,第二层以深度森林构建元模型,通过第一层的3个基模型分别获得反演结果P1, P2, P3,再将3个基模型的结果作为特征输入元模型,获得最终的海冰厚度反演结果P。

1.3.2 基模型和元模型

随机森林利用多棵树对样本进行训练和预测,是一种特殊的Bagging方法。它通过Bootstrap方法生成训练集,然后对每个训练集构造一棵决策树,并通过随机抽取特征中的最优解进行节点分裂,可以有效处理异源高维海冰数据,并充分利用不同数据集的优势特征组合,减少特征之间冗余带来的相互影响,进而防止过拟合。公式(1)中 T 为随机森林最终预测的结果, T_1, T_2, \dots, T_N 为 N 棵决策树预测的结果^[28]。

$$T = \frac{1}{N} \sum (T_1 + T_2 + \dots + T_N) \quad (1)$$

XGBoost由GBDT(Gradient boosting decision tree)改进而来,它将多个CART树模型通过串行集成的方式结合在一起,形成强学习器。XGBoost通过迭代训练回归树拟合非线性关系,利于挖掘特征与海冰厚度的复杂关联。XGBoost的目标函数如公式(2)所示,其中第一项为损失函数, y_i 为真实值, \hat{y}_i 为预测值,第二项为正则化用来防止过拟合^[29-30]。

$$\text{Obj}_k = \sum_{i=1}^M l(y_i, \hat{y}_i) + \Omega(f_k) \quad (2)$$

极端随机树是由GEURTS等提出^[31],是随机森林的变体,它在随机特征子集上构建决策树,无需特征选择,避免模型过度依赖特定特征,提升树间差异与多样性。该算法继承了随机森林的优势,可处理多维数据,降低对噪声敏感度,增强泛化能力^[32]。

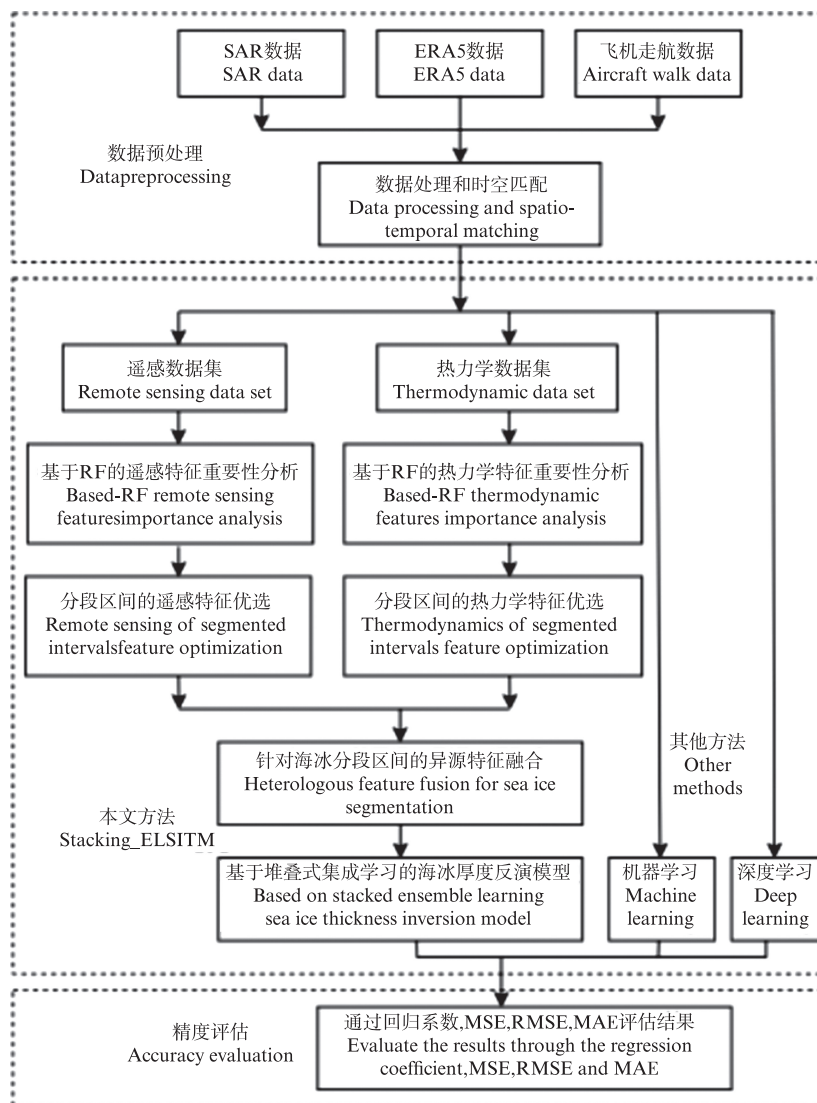


图6 基于堆叠式集成学习的海冰厚度反演总体框架

Fig. 6 Overall framework for sea ice thickness inversion based on stacked ensemble learning

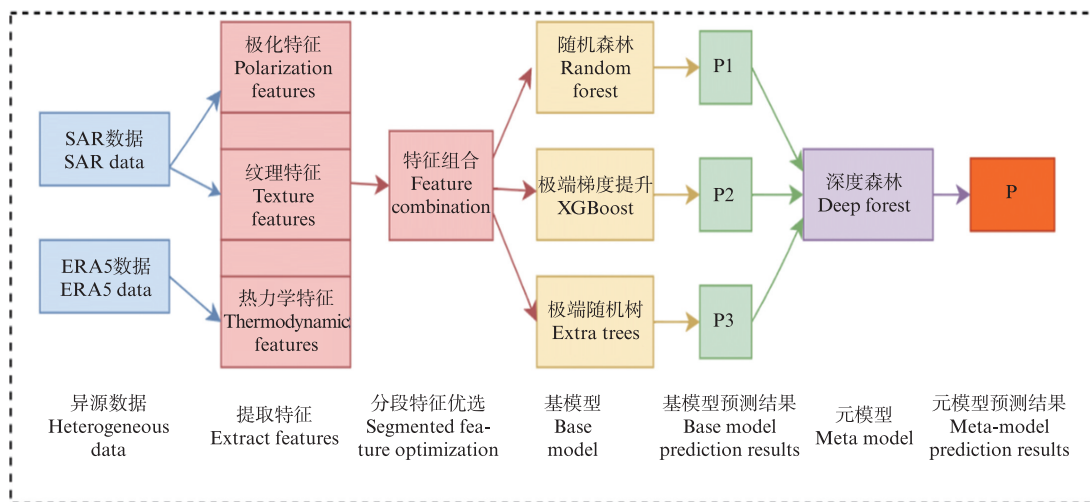


图7 堆叠式集成学习的海冰厚度反演模型

Fig. 7 Sea ice thickness inversion model based on stacking ensemble model

深度森林属于非神经网络式的深度模型,以多粒度级联森林构建。级联每层含2个随机森林(依基尼系数在随机特征子空间选分裂节点)和2个完全随机森林(全特征空间随机选特征分裂),多粒度扫描借类卷积操作扩维,强化特征关联学习,适配海冰特征优选与数据分段处理。其无需手动调参,达迭代条件自动停止,对整个建模流程采用交叉验证,以防止过拟合^[33-34]。

1.4 评价指标

海冰厚度反演属于回归问题,本文采用决定系数(R^2)、平均绝对误差(Mean square error, MAE)、均方误差(Mean squared error, MSE)和均方根误差(Root mean square error, RMSE)对反演结果进行评价,其中, R^2 代表预测值和真实值之间的拟合程度,其范围在(0,1)之间, R^2 越接近1,说明预测值和真实值之间差异越小,模型的反演效果越好;平均绝对误差(MAE)描述的是真实值与预测值差值的绝对值;均方误差(MSE)描述的是预测值和真实值差值的平方;均方根误差(RMSE)是在MSE的基础上再开平方,MAE、MSE、RMSE的误差值越小,代表模型的反演效果越好,评价指标的计算公式如下:

$$R^2 = 1 - \frac{\sum_i^n (y_i - \hat{y}_i)^2}{\sum_i^n (y_i - \bar{y})^2} \quad (3)$$

$$MAE = \sum_i^n \frac{|y_i - \hat{y}_i|}{n} \quad (4)$$

$$MSE = \frac{\sum_i^n (y_i - \hat{y}_i)^2}{n} \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_i^n (y_i - \hat{y}_i)^2}{n}} \quad (6)$$

式中: \hat{y}_i 为预测海冰厚度; y_i 为真实海冰厚度; n 为样本的数量。

2 结果与分析

2.1 遥感特征的重要性分析

特征的重要性是指各个特征对海冰厚度反演的相对贡献程度,本文选择随机森林算法来分析遥感特征在分段海冰厚度反演中的重要性。表3显示了分段海冰厚度反演中的遥感特征的重要性,海冰厚度的分段范围按照0到1.5 m、0到2 m、0到3 m、0到4 m、0到5 m进行划分,以下分别简称为(0,1.5)m、(0,2)m、(0,3)m、(0,4)m和(0,5)m。

从表3中可以看出,在(0,1.5)m区间最重要

的特征是入射角,其值为0.650,随着海冰的厚度增加,入射角的重要性在降低,比如在(0,2)m和(0,5)m入射角重要性为0.550和0.504,分别降低了10%和14.6%。入射角重要性的降低可能与不同厚度海冰的复介电常数有关,在一定条件下,入射角和复介电常数之间呈现正相关关系,即入射角随着复介电常数的增加而增大,同时,复介电常数又与海冰厚度有关,通常情况下,随着海冰厚度的增加,海冰的复介电常数也会相应减小,即复介电常数和海冰厚度呈负相关关系。因此入射角会随着海冰厚度的增加而减小,其重要性也会随之降低。

在(0,1.5)m区间的特征重要性分析中,后向散射系数(HH, HV)的重要性分别为0.020, 0.032,随着海冰厚度的增加,后向散射系数的重要性在提高。比如在(0,2)m和(0,5)m后向散射系数(HH, HV)重要性分别为0.048, 0.055和0.047, 0.069,分别提升了2.8%, 2.3%和2.7%, 3.7%,后向散射系数(HH, HV)与海冰的表面粗糙度有关,海冰厚度越大,粗糙度越大,其后向散射系数越大,因此随着海冰厚度的增加,后向散射系数会变大,其重要性也会增加。

另外,纹理特征的对比度、相异性、同质性、二阶矩、能量、最大概率、熵、均值、方差、相关性在(0,1.5)m区间的特征重要性依次为0.014, 0.010, 0.016, 0.007, 0.007, 0.012, 0.013, 0.054, 0.113, 0.045。随着海冰厚度的增加,部分纹理特征如对比度、相异性、同质性、二阶矩、能量、最大概率、熵、相关性的重要性在提高,而另一部分纹理特征如均值、方差的重要性在降低。纹理特征与海冰厚度的关系体现为海冰厚度越大,其表面的纹理越粗糙。因此随着海冰厚度的增加,大部分纹理特征的重要性都在提升,但也有部分纹理特征的重要性在降低,体现了环境、成分等多因素影响下海冰纹理特征变化的复杂性。

2.2 基于遥感特征的分段海冰厚度反演对比

本文使用Stacking_ELSITM模型对不同厚度区间的海冰进行实验。使用的遥感特征分为两组,F1和F2,其中F1包括入射角(α),后向散射系数(HH, HV),记为F1={ α , HH, HV};F2是在F1基础上添加提取的10个纹理特征(GLCM)做辅助特征,记为F2={F1, GLCM}。以下实验的结果都是取3次实验结果的平均值,如表4和表5所示。

表3 分段海冰厚度反演中的遥感特征重要性分析
Tab. 3 Important analysis of remote sensing features in segmented sea ice thickness inversion

遥感特征 Remote sensing features	(0,1.5)m	(0,2) m	(0,3) m	(0,4) m	(0,5) m
入射角 Incidence angle	0.650	0.550	0.484	0.496	0.504
后向散射系数 HH	0.020	0.048	0.050	0.046	0.047
后向散射系数 HV	0.032	0.055	0.064	0.064	0.069
对比度 Contrast ratio	0.014	0.037	0.044	0.043	0.038
相异性 Diversity	0.010	0.023	0.028	0.027	0.027
同质性 Homogeneity	0.016	0.037	0.046	0.047	0.047
二阶矩 Second moment	0.007	0.015	0.019	0.020	0.020
能量 Energy	0.007	0.015	0.019	0.020	0.020
最大概率 Maximum probability	0.012	0.027	0.031	0.030	0.028
熵 Entropy	0.013	0.033	0.039	0.037	0.038
均值 Mean value	0.054	0.045	0.049	0.050	0.053
方差 Variance	0.113	0.062	0.061	0.062	0.062
相关性 Correlation	0.045	0.046	0.059	0.052	0.047

表4 基于 F1 特征的分段海冰厚度反演精度
Tab. 4 Accuracy of segmented sea ice thickness inversion based on F1 features

精度评估 Accuracy assessment	(0,1.5) m	(0,2) m	(0,3) m	(0,4) m	(0,5) m
使用的特征 Features used	F1	F1	F1	F1	F1
R^2	0.847	0.631	0.572	0.597	0.596
MSE	0.015	0.049	0.136	0.237	0.317
RMSE	0.124	0.222	0.370	0.487	0.563
MAE	0.079	0.152	0.256	0.333	0.373

表5 基于 F2 特征的分段海冰厚度反演精度
Tab. 5 Accuracy of segmented sea ice thickness inversion based on F2 features

精度评估 Accuracy assessment	(0,1.5) m	(0,2) m	(0,3) m	(0,4) m	(0,5) m
使用的特征 Features used	F2	F2	F2	F2	F2
R^2	0.880	0.750	0.724	0.731	0.735
MSE	0.012	0.033	0.088	0.151	0.207
RMSE	0.110	0.183	0.270	0.389	0.456
MAE	0.073	0.127	0.210	0.273	0.310

从表4的实验结果可以看出:在仅使用遥感特征组合 F1 进行海冰厚度反演时,在(0,1.5)m 取得最好的反演精度,其结果是 R^2 为0.847,MSE 为0.015, RMSE 为0.124, MAE 为0.079。随着厚度范围的增加,其反演精度有下降趋势,在(0,2) m R^2 为0.631 降低了21.6%,MSE、RMSE 和 MAE 分别增大了3.4%、9.8%和7.3%,在(0,3) m R^2 为0.572 降低了27.5%,MSE、RMSE 和 MAE 分别增大了12.1%、24.6%和17.7%,在(0,4) m 和(0,5) m 范围也有类似的趋势, R^2 逐步降低,而 MSE、RMSE 和 MAE 逐步增大。由特征重要性分析可知,入射角的重要性会随着海冰厚度的增加而减少,后向散射系数会随着海冰厚度增加而增加,

但是由于入射角在海冰厚度反演中起着更重要的作用,所以随着厚度范围的增加,海冰厚度的反演精度总体上在下降。因此入射角(α)和后向散射系数(HH,HV)更适合进行一年冰及薄冰的厚度反演,对于超过1.5 m 的厚冰,需要融合更多的特征。

另外,从表5的实验结果可以看出,使用遥感特征组合 F2 进行海冰厚度反演时,相比 F1 组合,各个厚度区间反演精度都有提升,在(0,1.5) m 区间 R^2 提升了3.3%,MSE 降低了0.3%,RMSE 降低了1.4%,MAE 降低了0.6%,在(0,2) m、(0,3) m、(0,4) m 和(0,5) m 区间, R^2 分别提升了11.9%、15.2%、13.4%和13.9%,RMSE 分别降低了3.9%、

10%、9.8%和10.7%,MAE和MSE也有类似的效果。根据之前的特征重要性分析,纹理特征会随着海冰厚度增加,其重要性基本上都随之增加,说明在各个区间融入纹理特征都能有效改进海冰厚度反演精度;另外,融入纹理特征后,不同范围内的海冰厚度反演精度提升效果有所不同,随着海冰厚度增加,其反演精度的提升效果更加明显,这是由于海冰厚度与纹理特征有密切关系,当海冰的厚度越大,其纹理越粗糙,对厚度反演的影响越大;反之当海冰厚度越小,其纹理相对光滑,对厚度反演的影响就会逐渐减低,因此纹

理特征的融入对于厚冰反演有更好的提升效果。

2.3 热力学特征的重要性分析

为了分析热力学特征对于不同分段区间海冰厚度反演重要性,本文使用随机森林算法对特征重要性进行分析。从图8可以看出,(0,1.5)m区间内最重要的前3个特征是雪层温度、10 m高度的纬向风(u)、2 m气温;(0,2)m区间内最重要的前3个特征是海表温度、10 m高度的纬向风(v)、潜热辐射;(0,3)m到(0,5)m区间最重要的前3个特征均为感热辐射、10 m高度的纬向风(u)、潜热辐射。

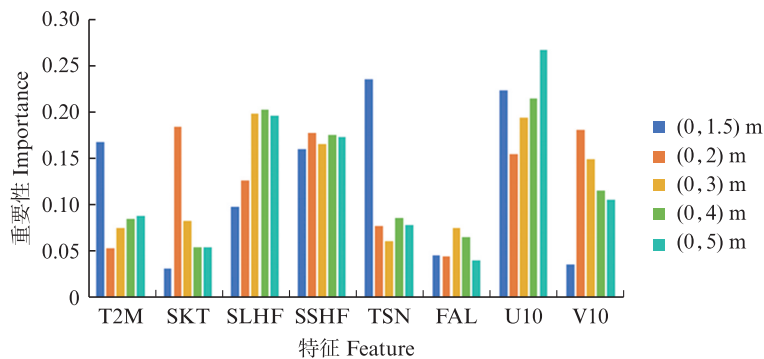


图8 热力学特征的重要性分析

Fig. 8 Importance analysis of the thermodynamic features

研究发现,北极海冰厚度变化与太阳辐射、风、温度、大气等多种因素相关,对于薄冰,温度起主导作用,在特征的重要性分析中,在(0,1.5)m的区间,2 m的气温和雪层温度对海冰厚度更加敏感,随着海冰厚度的增加,2 m气温和雪层温度的特征重要性开始有所下降,表明相比厚冰,薄冰对温度的变化更加敏感。对于厚冰,需要海冰和大气之间更多的热量交换才能使海冰厚度发生变化,在海冰与大气之间的相互作用下,上层海洋通过调节海洋与大气之间的温度差和湿度差向大气传输热量,同时大气通过风引起的潜热辐射变化来改变上层海洋的热量。从特征重要性分析也可以看出,随着海冰厚度的增加,在(0,3)m到(0,5)m区间,风速、感热辐射和潜热辐射的重要性最高。由于热力学参数之间存在较大的相关性,为了防止特征冗余增加模型的复杂度和过拟合等问题,后续实验中针对不同分段区

间分别选择最重要的前3个热力学特征进行融合。

2.4 基于遥感特征和热力学特征融合的海冰厚度反演

由于热力学特征在不同海冰厚度区间的特征重要性有所不同,因此本文分别使用遥感特征组合F2和针对不同海冰厚度区间的热力学优选特征进行融合以实现最优化的分段厚度反演。在(0,1.5)m区间选择热力学特征组合(雪层温度、10 m高度的纬向风(u)、2 m气温);(0,2)m区间选择热力学特征组合(海表温度、10 m高度的纬向风(v)、潜热辐射);(0,3)m到(0,5)m区间选择热力学特征组合(感热辐射、10 m高度的纬向风(u)、潜热辐射)分别构成优选特征组合F3^{opt}。基于遥感特征和热力学特征融合的分段海冰厚度反演结果如表6所示,不同分段区间的海冰厚度反演结果的散点图如图9所示。

表 6 基于遥感特征和热力学特征融合的分段海冰厚度反演结果
Tab. 6 Results of segmented sea ice thickness inversion based on the fusion of remote sensing features and thermodynamic features

精度评估 Accuracy assessment	(0,1.5) m	(0,2) m	(0,3) m	(0,4) m	(0,5) m
使用的特征 Features used	F3 ^{opt}	F3 ^{opt}	F3 ^{opt}	F3 ^{opt}	F3 ^{opt}
R^2	0.923	0.851	0.838	0.860	0.866
MSE	0.008	0.020	0.052	0.079	0.105
RMSE	0.089	0.142	0.228	0.281	0.324
MAE	0.054	0.092	0.149	0.179	0.199

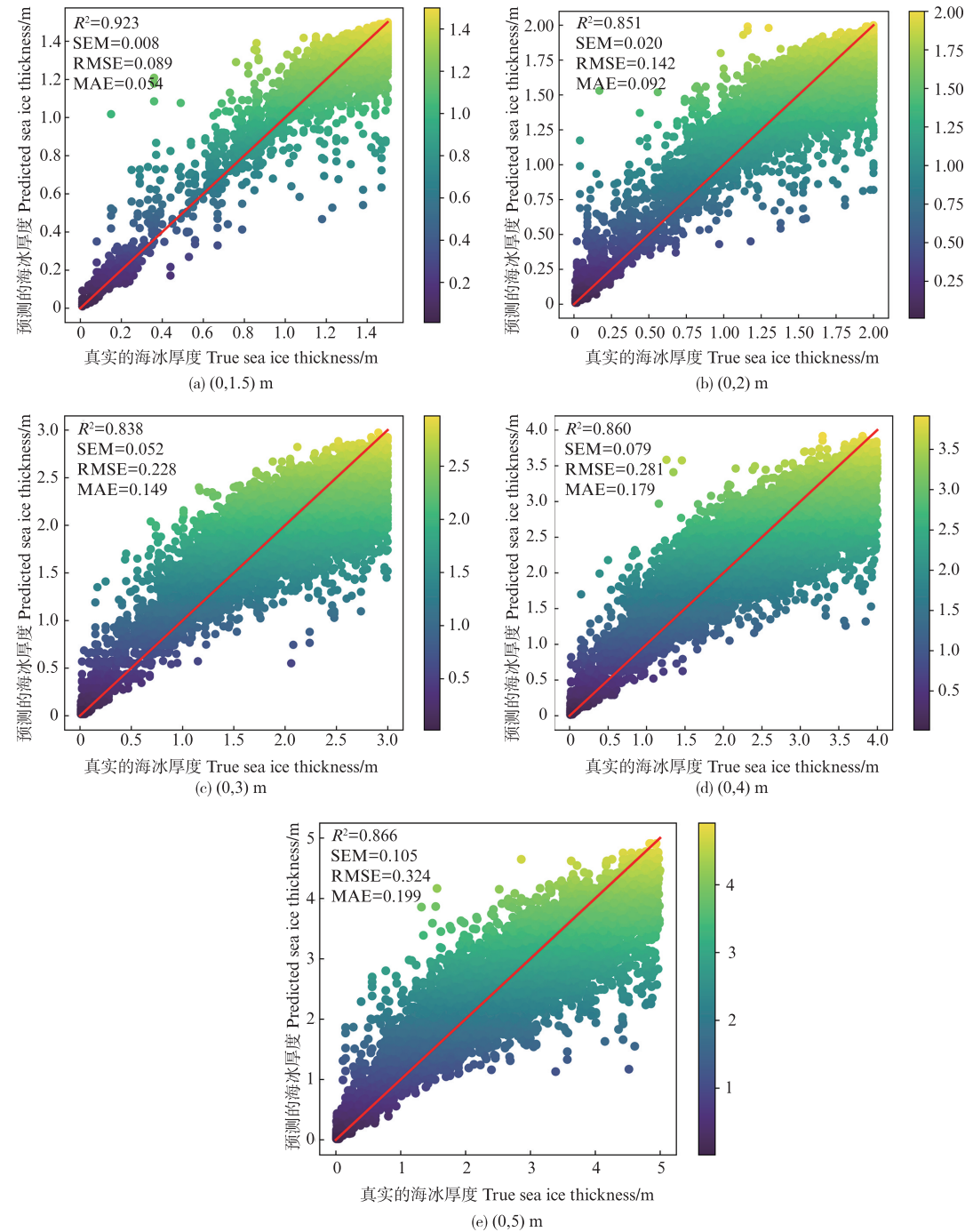


图 9 不同分段区间的海冰厚度反演散点图
Fig. 9 Scatter plot of sea ice thickness inversion in different segment intervals

从表6的实验结果可以看出,将遥感特征和热力学特征进行融合后,在(0,1.5) m区间仍然获得最高的反演精度, R^2 为0.923,MSE为0.008, RMSE为0.089,MAE为0.054。随着海冰厚度不断增大,反演精度总体上呈现下降趋势,在(0,5) m区间, R^2 为0.866,MSE为0.105, RMSE为0.324,MAE为0.199,说明随着海冰厚度的增加,海冰形成的原因和结构等复杂因素的影响加剧,一定程度上限制了反演精度的进一步提升。

另外,相比单独使用遥感特征,利用遥感特征和热力学特征融合进行海冰厚度反演,针对所有区间的海冰厚度反演精度均有明显提高,具体而言,在(0,1.5)m区间 R^2 提升了4.3%,RMSE降低了2.1%,(0,2)m区间 R^2 提升了10.1%,RMSE降低了4.1%,(0,3)m区间 R^2 提升了11.4%,RMSE降低了4.2%,(0,4)m区间 R^2 提升了12.9%,RMSE降低了10.8%,(0,5)m区间 R^2 提升了13.1%,RMSE降低了13.1%;同时可以看出,随着海冰厚度的增长,特征融合后的反演精度提升效果逐步提高,在薄冰(0,1.5)m区间,遥感特征已能够较为精准地捕捉海冰的相关信息,融合热力学特征后提升效果相对较低,而在厚冰区间,随着海冰厚度增加,其非线性特性更加复杂,利用优选的热力学特征进行异源数据分段特征融合后会进一步提升反演效果,尤其在(0,5)m厚度区间 R^2 提升最高达到13.1%,RMSE也降低最高达到13.1%。图9的海冰厚度反演散点图给出了异源数据融合后在不同分段区间的海冰厚度反演结果和真值之间的一致性展示。

2.5 基于不同模型的海冰厚度反演

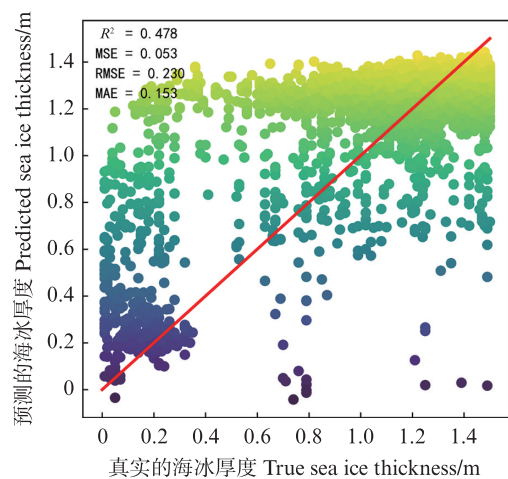
现有文献大多针对一年冰进行海冰厚度反演研究。为了与现有研究进行对比分析,本文利用遥感特征和热力学特征(TSN、U10、T2M)融合形成特征组合F3,对(0,1.5)m区间进行海冰厚度反演对比分析。对比方法包括支持向量回归(SVR)、梯度提升决策树(GDBT)、极端梯度提升(XGBoost)、深森林(DF)、随机森林(RF)、极端随机树(ERT)等6种机器学习方法,以及长短期记忆(LSTM)、门控递归单元(GRU)和卷积神经网络(CNN)等3种深度学习方法和2种深度学习的

组合方法 CNN-LSTM、CNN-GRU。本文中6种机器学习对比方法均采用了默认参数,5种深度学习模型均采用了最优化参数。所有对比方法以及本文提出的方法均基于相同的实验环境。均以3次实验的平均值作为最终结果,并根据精度评价指标和散点图对不同模型的海冰厚度预测结果进行对比分析和可视化评估。不同模型的对比分析结果见表7,散点图如图10所示。

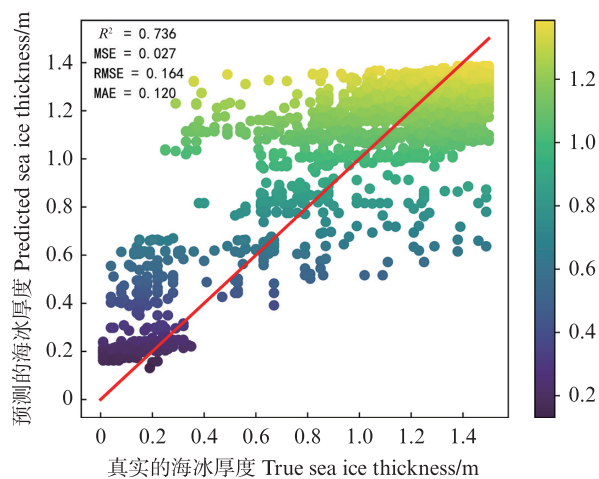
表7 基于不同模型的海冰厚度反演对比分析
Tab. 7 Accuracy of inversion of sea ice thickness for each model

模型名称 Model name	R^2	MSE	RMSE	MAE
SVR	0.478	0.053	0.230	0.153
GDBT	0.736	0.027	0.164	0.120
LSTM	0.771	0.023	0.153	0.095
GRU	0.751	0.025	0.159	0.102
CNN	0.809	0.019	0.139	0.090
XGBoost	0.879	0.012	0.111	0.076
DF	0.904	0.010	0.099	0.061
RF	0.913	0.009	0.094	0.057
ERT	0.917	0.008	0.092	0.055
CNN-LSTM	0.801	0.020	0.142	0.088
CNN-GRU	0.795	0.021	0.144	0.088
Stacking_ELSITM	0.923	0.008	0.089	0.054

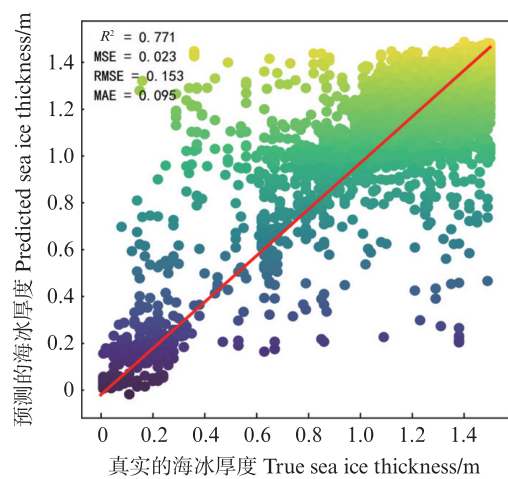
从表7可以看出,机器学习模型中,SVR模型难以有效拟合异源多特征与海冰厚度之间的复杂非线性关系,在海冰厚度反演中表现最差, R^2 仅为0.478,均方根误差为0.230;GDBT由于在处理非线性数据和避免过拟合方面的优势,相对SVR其反演精度有较明显提升;XGBoost结合了决策树的强大拟合能力和梯度提升的迭代优化策略,获得较好的反演精度, R^2 为0.879,均方根误差为0.111;DF通过多粒度级联结构捕捉海冰数据的局部和全局特征,明显提高了反演精度;RF通过集成多颗决策树的结果降低方差,提升了模型稳定性并减少了过拟合,获得较高的反演效果,ERT作为RF变体通过引入更多的随机性,进一步降低模型的方差,提升泛化能力,在海冰厚度反演中表现较突出, R^2 为0.917, RMSE为0.092。



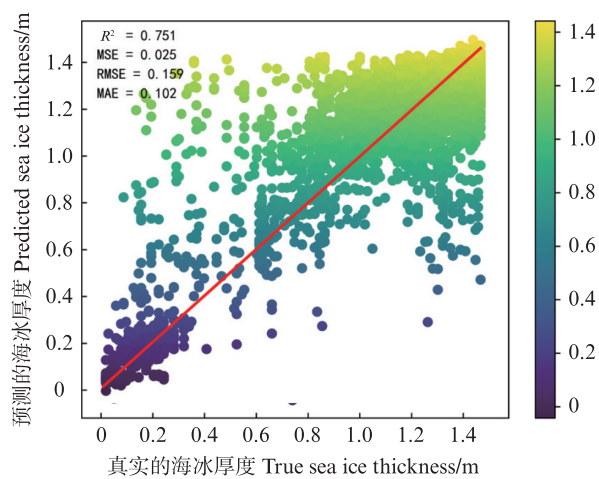
(a) SVR



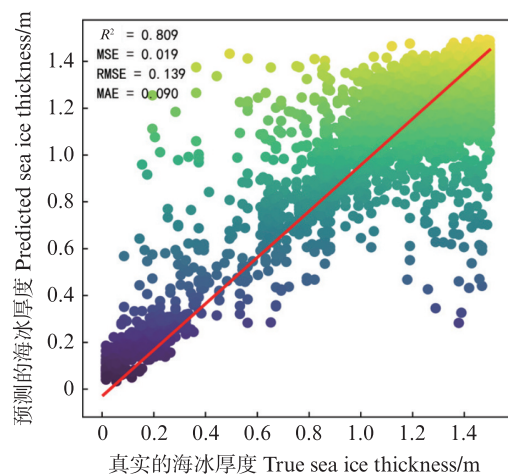
(b) GDBT



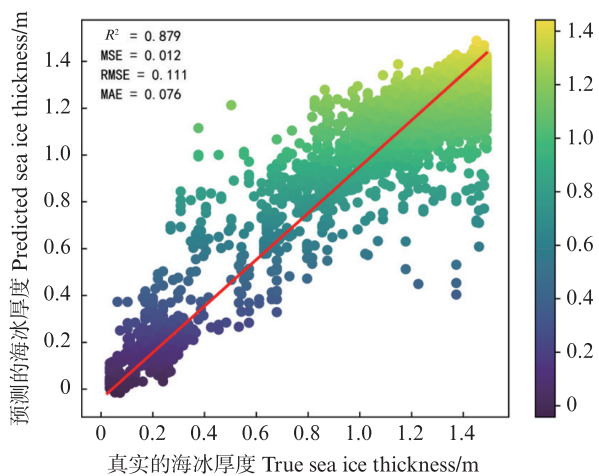
(c) LSTM



(d) GRU



(e) CNN



(f) XGBoost

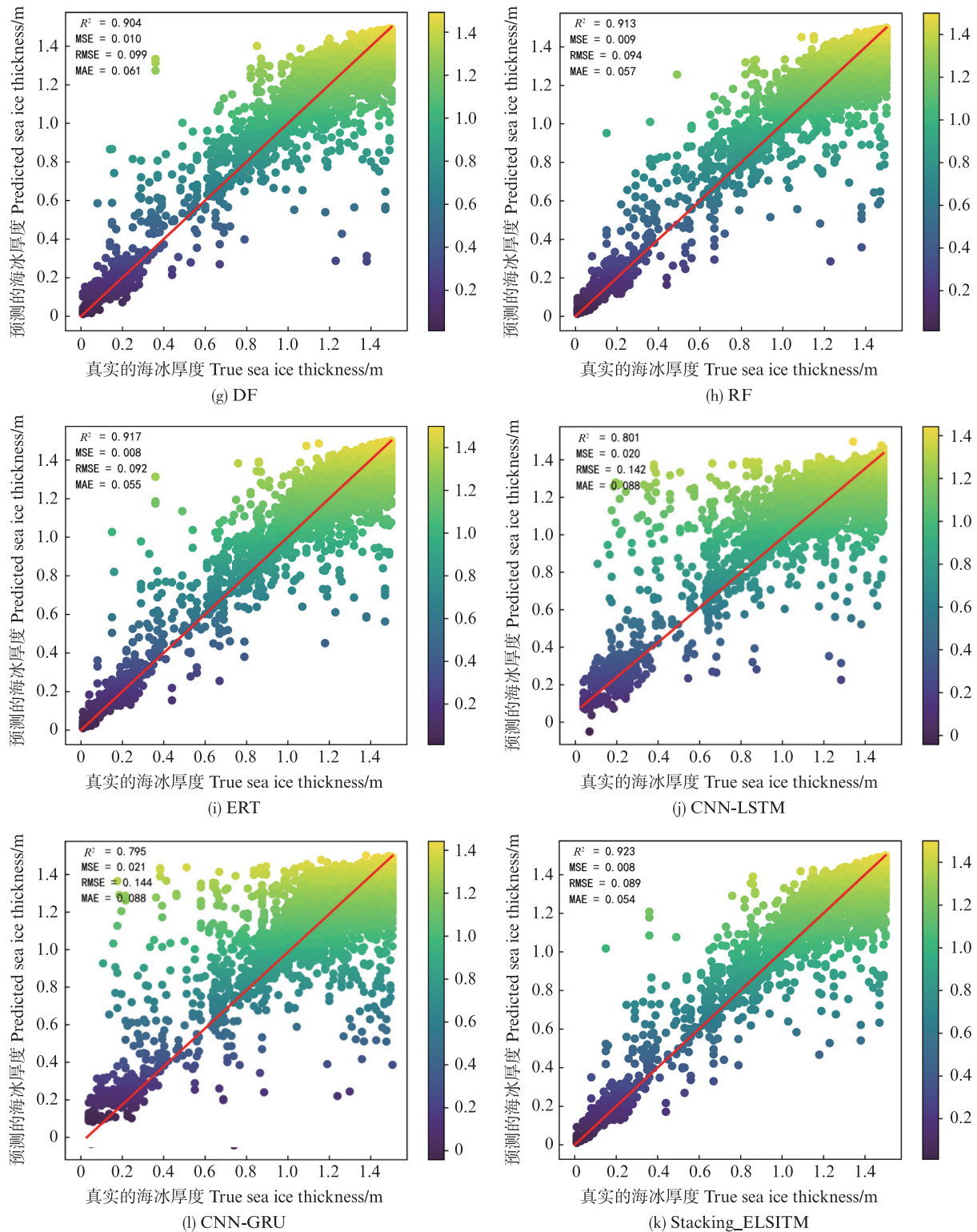


图 10 基于不同模型的海冰厚度反演散点图

Fig. 10 Scatter plot of sea ice thickness for each model

在基于深度学习的海冰厚度反演中,由于异源多特征与海冰厚度之间的复杂非线性关系和空间结构关联,3种深度学习模型 LSTM、GRU 和 CNN 模型通过增加 dropout 层来避免过拟合,一

定程度上提高了模型的性能, R^2 都能达到 0.75 以上,但整体上表现一般,与 GRU 和 LSTM 相比, CNN 擅长通过多维卷积核捕捉更多局部和空间细节信息,在 3 种模型中表现最好;两种深度学习

模型的组合总体上效果提升不明显。

本研究建议的 Stacking_ELSITM 方法由于利用随机森林、XGBoost、极端随机树和深度森林的优势构建堆叠集成学习模型,能充分利用异源数据特征与海冰厚度之间的隐藏关联,实现优势互补,获得了最好的反演效果。 R^2 达到 0.923,同时 RMSE 降低至 0.089。从图 10 的散点图也可以看出,相比其他方法,本研究建议的 Stacking_ELSITM 方法的散点图分布更为集中,更接近对角线($y=x$)的红线,说明其预测值和真实值的一致性更好,即模型反演精度更高。

3 总结

本文利用哨兵一号 SAR 数据和 ERA5 热力学数据进行异源数据融合,提出基于堆叠式集成学习(Stacking_ELSITM)分段特征融合的海冰厚度反演模型,通过异源数据分段特征优选和特征融合,获得更加精确的海冰厚度反演结果,具体总结如下:

(1)针对单一 SAR 数据的海冰厚度反演中,仅使用入射角和后向散射系数(HH, HV)在(0, 1.5)m 区间获得最好的反演精度,随着厚度范围增加,海冰厚度的反演精度总体上有所下降;融入纹理特征后,不同区间内的海冰厚度反演精度都有提升,但由于海冰厚度越大,纹理越粗糙,纹理特征对厚度反演的影响也会越大,因此随着海冰厚度增加,反演精度的提升效果也更加明显。

(2)针对异源数据的海冰厚度反演中,通过 SAR 数据特征和热力学特征的融合,在所有的厚度区间上都获得比使用单源 SAR 数据更好的反演结果,而且随着海冰厚度的增长,影响因素的复杂性导致单源遥感特征的局限性更加明显,通过分段优选热力学特征的有效融合,厚冰区间范围的反演精度提升效果也更加明显。

(3)无论利用单源数据还是异源数据融合,(0, 1.5)m 区间始终能获得最佳的反演效果,随着海冰厚度区间的扩大,反演精度都有不同程度的下降,通过异源数据特征优选和分段特征融合,可以提高各个分段区间的海冰反演精度,进而获得更好的整体反演效果。

(4)与其他方法相比,本文建议的 Stacking_ELSITM 通过对不同分段范围的海冰进行特征优选组合,利用多个基模型和元模型之间的串并级

联实现优势互补,充分挖掘了异源数据特征与海冰厚度之间的隐藏关联,通过堆叠式集成学习的方式,实现了对海冰厚度的精确反演,获得了最好的反演效果。

精确的海冰厚度反演对气候变化监测、北极航道通行和海洋利用具有重要的应用价值。本文借助遥感技术对极地海冰进行分段特征融合,实现了不同区间海冰厚度的高精度反演,为极地海冰探测提供了新的思路和方法。此外,鉴于积雪和冰融水对雷达信号的传输有一定的影响,后续我们将进一步考虑积雪和冰融水对海冰反演的影响。

作者声明本文无利益冲突。

参考文献:

- [1] 刘骥平, 雷瑞波, 宋米荣, 等. 适应极地快速变化海冰模式的研发与挑战[J]. 大气科学学报, 2021, 44(1): 12-25.
LIU J P, LEI R B, SONG G R, et al. Development and challenge of sea ice model adapting to rapid polar sea ice changes [J]. Transactions of Atmospheric Sciences, 2021, 44(1): 12-25.
- [2] 张建, 杨元德, 杨全明, 等. 基于 IceBridge 数据的南极别林斯高晋海的海冰厚度研究[J]. 冰川冻土, 2020, 42(2): 295-306.
ZHANG J, YANG YD, YANG QM, et al. Sea ice thickness research in the Bellingshausen Sea from IceBridge [J]. Journal of Glaciology and Geocryology, 2020, 42(2): 295-306.
- [3] 史茜. 基于深度学习的同异源遥感影像海冰分类[D]. 上海: 上海海洋大学, 2022.
SHI X. Homologous and heterologous remote sensing seaice classification based on deep learning [D]. Shanghai: Shanghai Ocean University, 2022.
- [4] 王昀, 李雪薇, 王今菲, 等. 北极海冰厚度的热力学和动力学影响因素研究评述[J]. 极地研究, 2023, 35(1): 124-138.
WANG Y, LI X W, WANG J F, et al. Overview on the thermodynamic and dynamic factors influencing Arctic sea ice thickness [J]. Chinese Journal of Polar Research, 2023, 35(1): 124-138.
- [5] 季青, 庞小平, 许苏清, 等. 极地海冰厚度探测方法及其应用研究综述[J]. 极地研究, 2016, 28(4): 431-441.
JI Q, PANG X P, XU S Q, et al. Review of technology and application research on polar sea ice thickness detection [J]. Chinese Journal of Polar Research, 2016,

- 28(4): 431-441.
- [6] 罗丽程, 张文奇, 胡勇, 等. 基于 Sentinel-1 卫星的北冰洋海冰信息提取[J]. 上海航天, 2018, 35(3): 16-22.
- LUO L C, ZHANG W Q, HU Y, et al. Information extraction of Arctic sea ice based on Sentinel-1 Satellite [J]. Aerospace Shanghai, 2018, 35(3): 16-22.
- [7] 李明慧. 基于深度学习的 SAR 影像海冰分类研究[D]. 上海: 上海海洋大学, 2019.
- LI M H. Sea ice classification based on deeplearning with SAR imagery [D]. Shanghai: Shanghai Ocean University, 2019.
- [8] 葛梦滢, 高稳, 祝敏, 等. 基于 SE-ConvLSTM 的时空特征融合 SAR 图像海冰分类[J]. 遥感技术与应用, 2023, 38(6): 1306-1316.
- GE M Y, GAO W, ZHU M, et al. Sea ice classification of SAR images based on SE-ConvLSTM spatial-temporal feature fusion [J]. Remote Sensing Technology and Application, 2023, 38(6): 1306-1316.
- [9] 沈航. 基于深度学习的异源遥感影像多特征融合海冰分类方法研究[D]. 上海: 上海海洋大学, 2023.
- SHEN H. Research on multi-feature fusion sea ice classification method of heterogeneous remotesensing images based on deep learning [D]. Shanghai: Shanghai Ocean University, 2023.
- [10] WAKABAYASHI H, MATSUOKA T, NAKAMURA K, et al. Polarimetric characteristics of sea ice in the Sea of Okhotsk observed by airborne L-band SAR [J]. IEEE Transactions on Geoscience and Remote Sensing, 2004, 42(11): 2412-2425.
- [11] KIM J W, KIM D J, HWANG B J. Characterization of Arctic sea ice thickness using high-resolution spaceborne polarimetric SAR data [J]. IEEE Transactions on Geoscience and Remote Sensing, 2012, 50(1): 13-22.
- [12] ZHANG X, ZHANG J, MENG J M, et al. Analysis of multi-dimensional SAR for determining the thickness of thin sea ice in the Bohai Sea [J]. Chinese Journal of Oceanology and Limnology, 2013, 31(3): 681-698.
- [13] 刘眉洁, 戴永寿, 张杰, 等. 拉布拉多海一年平整冰厚度 SAR 反演算法[J]. 中国石油大学学报(自然科学版), 2014, 38(3): 186-192.
- LIU M J, DAI Y S, ZHANG J, et al. First-year level sea-ice thickness retrieval in Labrador Sea using C-band polarimetric SAR data [J]. Journal of China University of Petroleum (Edition of Natural Science), 2014, 38(3): 186-192.
- [14] 刘立剑. 基于 C 波段 PolSAR 的多特征融合的海冰厚度反演[D]. 大连: 大连海事大学, 2015.
- LIU L J. Sea-Ice thickness inversion of multi-feature fusion based on C-band PolSAR images [D]. Dalian: Dalian Maritime University, 2015.
- [15] 于淼, 卢鹏, 李志军, 等. 基于 SAR 图像纹理的北极海冰厚度的反演研究[J]. 极地研究, 2018, 30(3): 329-337.
- YU M, LU P, LI Z J, et al. Arctic sea ice thickness retrieval based on SAR image texture feature [J]. Chinese Journal of Polar Research, 2018, 30(3): 329-337.
- [16] HERBERT C, MUNOZ-MARTIN J F, LLAVERIA D, et al. Sea ice thickness estimation based on regression neural networks using L-band microwave radiometry data from the FSSCat mission [J]. Remote Sensing, 2021, 13(7): 1366.
- [17] SHAMSHIRI R, EIDE E, HØYLAND K V. Spatio-temporal distribution of sea-ice thickness using a machine learning approach with Google Earth Engine and Sentinel-1 GRD data [J]. Remote Sensing of Environment, 2022, 270: 112851.
- [18] LIANG Z Y, JI Q, PANG X P, et al. Estimation of daily arctic winter sea ice thickness from thermodynamic parameters using a self-attention convolutional neural network [J]. Remote Sensing, 2023, 15(7): 1887.
- [19] ZHAO L, XIE T, PERRIE W, et al. Deep-learning-based sea ice classification with sentinel-1 and AMSR-2 data [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16: 5514-5525.
- [20] ZHAO P, HE Z B. Evaluation of ERA5 reanalysis temperature data over the Qilian Mountains of China [J]. Journal of Mountain Science, 2025, 22(1): 198-209.
- [21] 于淼. 利用 RadarSat-2 SAR 影像反演北极夏季海冰厚度的研究[D]. 大连: 大连理工大学, 2018.
- YU M. Study on sea ice thickness retrieval in Arctic summer using RadarSat-2 SAR imagery [D]. Dalian: Dalian University of Technology, 2018.
- [22] HENDRICKS S, RICKER R, HAAS C, et al. Airborne sea ice plus snow thickness during the PAMARCMIP 2017 aircraft campaign in the Arctic Ocean [J]. PANGAEA, 2020.
- [23] 韩彦岭, 高仪, 王静, 等. 结合未标签样本和 CNN 的高光谱遥感图像分类[J]. 遥感信息, 2020, 35(5): 19-30.
- HAN Y L, GAO Y, WANG J, et al. Combining unlabeled samples and CNN for hyperspectral remote sensing image classification [J]. Remote Sensing Information, 2020, 35(5): 19-30.
- [24] 孔毅, 陈曦, 艾未华, 等. 基于 GLCM 和小波特征的海冰分类[J]. 解放军理工大学学报(自然科学版), 2015, 16(1): 74-79.
- KONG Y, CHEN X, AI W H, et al. SAR ice classification based on GLCM and wavelet feature [J]. Journal of Army Engineering University of PLA, 2015, 16(1): 74-79.

- [25] 张晰, 张杰, 纪永刚. 基于纹理特征分析的辽东湾 SAR 影像海冰检测[J]. 海洋科学进展, 2008, 26(3): 386-393.
ZHANG X, ZHANG J, JI Y G. SAR imaging sea ice detection in Liaodong Bay based on texture feature analysis[J]. Advances in Marine Science, 2008, 26(3): 386-393.
- [26] 岳瀚栋, 窦挺峰, 李润奎, 等. 基于多种机器学习及其堆叠式集成方法的月尺度北极海冰预测研究[J]. 冰川冻土, 2023, 45(3): 893-901.
YUE H D, DOU T F, LI R K, et al. Monthly-scale Arctic sea ice extent prediction based on multiple machine learning and stacking ensemble methods[J]. Journal of Glaciology and Geocryology, 2023, 45(3): 893-901.
- [27] 江乾坤, 王成哲. 互联网企业海外并购财务风险大数据预警研究——基于 Stacking 集成学习[J]. 技术经济, 2023, 42(9): 147-160.
JIANG Q K, WANG C Z. Research on big data early warning of financial risks of overseas mergers and acquisitions of internet enterprises based on Stacking integrated learning[J]. Journal of Technology Economics, 2023, 42(9): 147-160.
- [28] 邓红卫, 罗亮. 基于 SMA 算法优化随机森林的 PPV 预测模型[J]. 黄金科学技术, 2023, 31(4): 624-634.
DENG H W, LUO L. PPV prediction model based on random forest optimized by SMA algorithm[J]. Gold Science and Technology, 2023, 31(4): 624-634.
- [29] 唐钦锡, 王涛. 基于 XGBoost 的压裂水平井产能预测[J]. 中国石油和化工标准与质量, 2023, 43(24): 15-17.
TANG Q X, WANG T. Capacity prediction of fracturing horizontal wells based on XGBoost[J]. China Petroleum and Chemical Standard and Quality, 2023, 43(24): 15-17.
- [30] 赵兵朝, 张晴, 王京滨, 等. 基于 SSA-XGBoost 模型的地表下沉系数预测研究[J]. 矿业研究与开发, 2024, 44(2): 89-95.
ZHAO B C, ZHANG Q, WANG J B, et al. Study on prediction of surface subsidence coefficient based on SSA-XGBoost model[J]. Mining Research and Development, 2024, 44(2): 89-95.
- [31] 饶骁驰, 杨昊, 喻辉, 等. 基于极端随机树算法的流体识别研究[J]. 物探化探计算技术, 2023, 45(5): 566-578.
RAO X C, YANG H, YU H, et al. Research on fluid recognition with extremely randomized trees algorithm[J]. Computing Techniques for Geophysical and Geochemical Exploration, 2023, 45(5): 566-578.
- [32] 于海洋, 谢赛飞, 郭灵辉, 等. 融合光谱和空间特征的土壤重金属含量极端随机树估算[J]. 农业机械学报, 2022, 53(8): 231-239.
YU H Y, XIE S F, GUO L H, et al. Extremely randomized trees estimation of soil heavy metal content by fusing spectra and spatial features[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8): 231-239.
- [33] ZHOU Z H, FENG J. Deep forest[J]. National Science Review, 2019, 6(1): 74-86.
- [34] 胡小晨, 郭宁, 沈拓, 等. 基于主成分分析和深度森林算法的 S700K 转辙机故障诊断[J]. 同济大学学报(自然科学版), 2024, 52(1): 35-40.
HU X C, GUO N, SHEN T, et al. PCA-gcForest-based fault diagnosis of S700K switch machine[J]. Journal of Tongji University (Natural Science), 2024, 52(1): 35-40.

Sea ice thickness inversion method based on segmented filtering feature fusion of heterologous remote sensing data

JIA Piaopiao¹, HAN Yanling¹, HE Haiyang², WANG Jing¹, YANG Shuhu¹, ZHANG Yun¹, HONG Zhonghua¹

(1. College of Information, Shanghai Ocean University, Shanghai 201306, China; 2. Shanghai SenseTime Intelligent Technology Co., Ltd., Shanghai 201306, China)

Abstract: To address the issue that the reliance on a single data source in current sea ice thickness detection limits the further improvement of sea ice thickness inversion accuracy. This paper proposes a sea ice thickness inversion method based on segmented feature fusion of heterogeneous data, the experiment uses Sentinel-1 synthetic aperture radar (SAR) data and ERA5 reanalysis data, by dividing sea ice thickness intervals (e.g., (0, 1.5) m, (0, 2) m, (0, 3) m), optimal feature combinations are selected for sea ice in different segmented intervals, a sea ice thickness inversion model based on stacked ensemble learning is constructed, which realizes complementary advantages through the series-parallel cascade of multiple base models and meta-models, this approach fully explores the hidden correlations between heterogeneous features and sea ice thickness to achieve accurate inversion of segmented sea ice thickness. The results show that compared with other traditional machine learning methods, this method achieves better overall inversion performance across different segmented intervals, notably, the interval of (0, 1.5) m exhibits the best performance, with a coefficient of determination (R^2) reaching 0.923 and a root mean square error (RMSE) as low as 0.089 m. The study demonstrates that segmented feature optimization and heterogeneous data fusion can effectively improve the inversion accuracy of sea ice thickness, verifying the advantages of the proposed stacked ensemble learning model in heterogeneous data fusion. This research provides a new method for achieving high-precision inversion of sea ice thickness.

Key words: sea ice thickness inversion; heterogeneous data fusion; segmented filtering feature fusion; stacking ensemble learning; Arctic