doi:10.6041/j.issn.1000-1298.2026.17.005

基于 Sentinel－1/2 时序数据与历史信息融合的甘蔗分布提取方法

孙琬婷[1]　何　芸[2]　朱秀芳[1]　李　乐[3]

（1.北京师范大学遥感与数字地球全国重点实验室，北京 100875；2.自然资源部国土卫星遥感应用中心，北京 100048；
3.广东工业大学管理学院，广州 510520）

摘要：精准提取甘蔗种植分布对我国糖料作物监测和食糖供应安全评估具有重要意义。针对现有方法以单年遥感观测为主、未能充分利用甘蔗多年连续种植先验信息的问题，本研究以广西崇左市为研究区，基于 Sentinel－2 L2A 与 Sentinel－1 SAR 时序数据，构建了融合多年历史信息的长短期记忆网络（LSTM）多分支分类框架，系统对比了历史数据（前 n 年原始遥感特征）与历史概率（前 n 年基线模型输出的连续甘蔗归属概率）两种先验信息引入策略在 2022—2025 年 4 个目标年份中的分类效果。结果表明，引入 3 年历史信息后，各数据源方案的分类精度均显著提升。其中，“Sentinel－1/2 融合＋3 年历史数据”方案表现最优，其在 4 年平均条件下，F1 值达到 90.6%、Kappa 系数为 0.854；相较于同源单年基线模型，该方案的 F1 值提升约 8.3 个百分点；在 3 年回溯期条件下，历史数据与历史概率两种策略的 F1 值均稳定在 90% 以上，表现出较高的一致性；随着回溯年限由 1 年延长至 3 年，模型分类精度呈整体上升趋势，该最佳回溯年限与广西甘蔗“一年新植、两年宿根”的 3 年农艺种植周期高度吻合。此外，Sentinel－1 与 Sentinel－2 的融合在基线场景下未能展现显著协同效应，但在引入 3 年历史信息后，融合方案精度明显优于仅光学方案，表明 SAR 数据的结构敏感性在时序先验信息的支持下能够充分发挥作用。基于最优方案，本研究生成了崇左市 2022—2025 年 10 m 分辨率甘蔗分布结果，验证了该方法在多年连续制图中的可行性。

关键词：甘蔗；分布提取；长短期记忆网络；Sentinel－1；Sentinel－2；时序分类

中图分类号：S127　　**文献标识码**：A　　**文章编号**：1000-1298(2026)17-0054-11　　**OSID**：

Sugarcane Distribution Extraction Method Based on Sentinel－1/2 Time-series Data and Historical Information Fusion

Sun Wanting[1]　He Yun[2]　Zhu Xiufang[1]　Li Le[3]

(1. *State Key Laboratory of Remote Sensing and Digital Earth*, *Beijing Normal University*, *Beijing* 100875, *China*
2. *Land Satellite Remote Sensing Application Center*, *Ministry of Natural Resources*, *Beijing* 100048, *China*
3. *School of Management*, *Guangdong University of Technology*, *Guangzhou* 510520, *China*)

Abstract: Accurately mapping sugarcane planting distribution is of great significance for monitoring the sugar crop industry and assessing sugar supply security in China. To address the limitation of existing methods that primarily relied on single-year remote sensing observations and failed to fully utilize multi-year continuous planting prior information, the Chongzuo City in Guangxi Zhuang Autonomous Region was selected as the study area. A multi-branch long short-term memory (LSTM) classification framework integrating multi-year historical information was constructed based on time-series data from Sentinel－2 L2A and Sentinel－1 SAR imagery. Two strategies for incorporating prior information were systematically compared: the historical data strategy (using raw remote sensing features from the previous n years) and the historical probability strategy (using continuous sugarcane probability maps output by a baseline model for the previous n years). The classification performance was evaluated for four target years from 2022 to 2025. The results demonstrated that introducing three years of historical information significantly improved classification accuracy across all data source schemes. Among all configurations, the fusion of Sentinel－1 and Sentinel－2 combined with three years of historical data achieved the best performance,

收稿日期：2026－05－20　修回日期：2026－05－26

基金项目：广东省基础与应用基础研究基金项目（2023A1515010897）

作者简介：孙琬婷（2001—），女，硕士生，主要从事农业遥感研究，E-mail：202311079929@mail.bnu.edu.cn

通信作者：朱秀芳（1982—），女，教授，博士生导师，主要从事农业遥感研究，E-mail：zhuxiufang@bnu.edu.cn

with an average F1-score of 90.6%, a Kappa coefficient of 0.854, representing an improvement of approximately 8.3 percentage points in F1-score compared with the single-year baseline model. Under the three-year retrospective condition, both the historical data and historical probability strategies yielded F1-scores consistently above 90%, showing high consistency between the two approaches. As the retrospective period increased from one to three years, the classification accuracy exhibited an overall upward trend, and this optimal retrospective period of three years coincided well with the typical three-year agronomic cycle of sugarcane cultivation in Guangxi, characterized as "one year of newly planted cane followed by two years of ratoon cane". Furthermore, the fusion of Sentinel－1 and Sentinel－2 data did not demonstrate a significant synergistic effect in the baseline scenario without historical information; however, after incorporating three years of historical information, the fusion scheme significantly outperformed the optical-only scheme, indicating that the structural sensitivity of SAR data can be fully leveraged with the support of temporal prior information. Based on the optimal scheme, it generated 10-meter resolution sugarcane distribution maps for Chongzuo City from 2022 to 2025, confirming the feasibility and effectiveness of the proposed method for multi-year continuous sugarcane mapping.

Key words: sugarcane; distribution extraction; long short-term memory network; Sentinel－1; Sentinel－2; time-series classification

0　引言

甘蔗是全球重要的糖料作物和生物能源作物，为全球 80% 的食糖生产提供原料[1]。中国是全球第三大甘蔗生产国，广西壮族自治区作为中国最大的甘蔗主产区，贡献了全国甘蔗产量的 60% 以上[2-3]。崇左市位于广西西南部，是广西甘蔗种植面积和产量均居首位的核心产区，甘蔗产业对当地经济具有支柱性意义。准确、及时地获取甘蔗种植区空间分布信息，对于食糖供应安全评估、农业政策制定以及产业结构优化具有重要意义[4-5]。

遥感技术因其覆盖范围广、重访周期短的优势，已成为大范围作物制图的主要手段[6-8]。然而，甘蔗的遥感提取面临多重挑战。首先，甘蔗与玉米、高粱等 C4 作物在物候特征上高度相似，光谱混淆现象突出[9]。其次，甘蔗种植具有宿根特性，同一地块可连续多年生长，但也存在轮作现象，这使得年际间的种植分布呈现复杂的动态变化[10-11]。因此，如何充分利用多源遥感数据和时序信息，提高甘蔗分类的准确性和稳定性，一直是研究的热点。

近年来，深度学习方法，特别是长短期记忆网络（Long short-term memory, LSTM），因其能够捕捉作物时序物候特征，在作物制图中展现出优越性能[12]。LSTM 能够有效处理遥感时间序列数据，学习作物的生长规律，从而提升分类精度。在数据源层面，Sentinel－2 多光谱影像和 Sentinel－1 雷达影像因其免费获取、时空分辨率高的优势，已成为作物制图的主流数据源[13-14]。Sentinel－2 提供丰富的光谱信息，能够反映作物的绿度变化和物候特征[15]；Sentinel－1 则对作物结构和水分敏感，可弥补光学影像受云雨天气影响的不足[16]。然而，光学与雷达数据在甘蔗分类中的融合效果尚存在争议，多源融合是否优于单一数据源，缺乏系统的实证评估。

历史信息的利用是作物分布制图中的一个重要方向。由于甘蔗种植具有宿根特性，同一地块可连续多年生长，引入历史样本信息可以提升当年作物的识别能力[17-18]。现有研究多采用将历史遥感数据作为额外输入特征的方式[19-20]，或是通过时序模型捕捉多年变化规律[21]。然而，历史信息的引入也带来了新的问题：历史信息究竟应以原始遥感数据的形式输入，还是应采用历史年份的分类结果（即作物归属概率），尚无定论；历史数据的最佳回溯年限有待确定；不同回溯条件下历史信息的作用是否一致也缺乏验证。这些问题直接关系到作物制图方法的选择和推广，但目前尚缺乏系统性的评估。特别是在甘蔗这一具有宿根特性的作物上，历史信息利用策略的研究几乎空白。

针对上述研究缺口，本文以广西壮族自治区崇左市为研究区，基于 Sentinel－1/2 时序数据，构建面向多年历史信息融合的多分支 LSTM 分类框架，设计 14 组对比实验，系统开展以下研究：评估引入多年历史信息对基于 Sentinel 数据的甘蔗分类精度的提升效果；对比历史数据与历史概率两种引入方式在相同回溯年限下的性能差异，并剖析其内在机理；探究历史信息的最佳回溯年限及其与甘蔗农艺周期的对应关系。

1　研究区及数据

1.1　研究区概况

本研究以广西壮族自治区崇左市为研究区

(图 1)。崇左市位于广西西南部(21°36′~23°22′N,106°33′~108°06′E),总面积约 1.73×10^4 km^2。该市下辖江州区、扶绥县、宁明县、龙州县、大新县、天等县和凭祥市共 7 个县区,其中江州区、扶绥县、龙州县及大新县东部和宁明县北部为甘蔗主要种植区[22]。研究区属亚热带季风气候,水热资源丰富,是广西甘蔗种植的核心区域。地形以低山丘陵为主,种植结构相对单一,甘蔗连片种植特征明显,是开展甘蔗遥感识别方法研究的理想区域[23]。

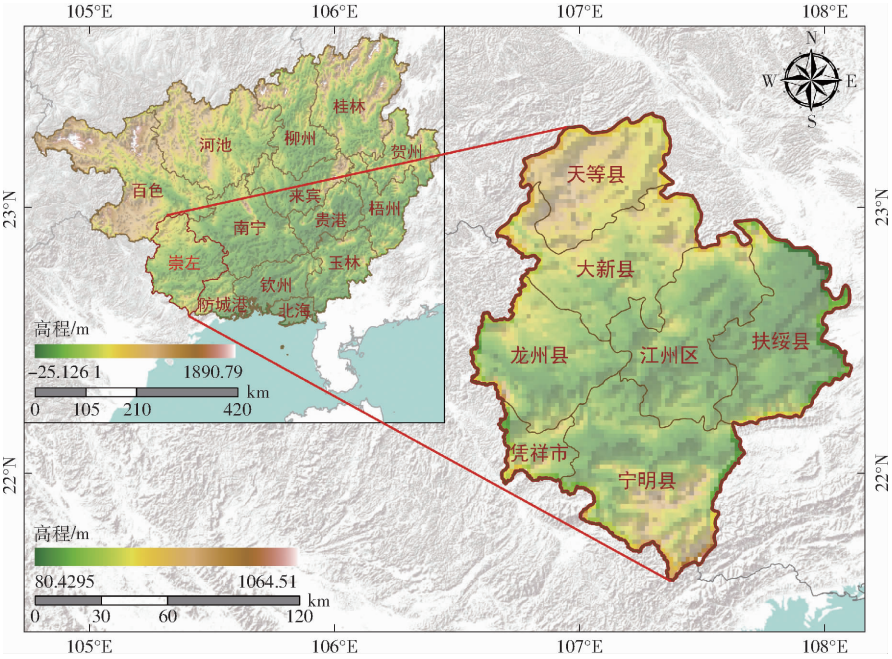

图 1 研究区域

Fig. 1 Study area

1.2 数据获取与预处理

1.2.1 Sentinel-2 L2A 数据

本研究使用 Sentinel-2 L2A 地表反射率产品,考虑到 2018 年及以前崇左地区该产品的月度有效覆盖数量不足,数据获取时间范围限定为 2019—2025 年。为减少云层干扰,采用 Cloud Score + 产品[24]进行云掩膜,将 cs_cdf 不低于 0.60 的像元判定为清晰像元。在此基础上,将全年影像按月进行中位数合成,生成 12 幅月度影像。针对月度合成后仍存在数据空洞的像元,首先采用时间序列线性插值进行填补,以前后相邻有效月份的加权平均估计缺失月份的反射率,该方式能够较好地保留作物物候的季节性演化趋势。若某像元在年内存在连续多月缺失导致插值失效,则以该像元全年有效观测的中位数进行回补,避免训练样本中出现完全无效的像元,确保时间序列的完整性与连续性。

为进一步消除时序中的高频噪声并保留地物物候变化的季节性趋势,对月度反射率序列应用 Savitzky-Golay 滤波[25],设置窗口大小为 5,采用二次多项式拟合进行平滑。基于平滑后的反射率,计算 6 个常用遥感指数:归一化差值植被指数(Normalized difference vegetation index, NDVI)[26]、归一化差值水体指数(Normalized difference water index, NDWI)[27]、增强型植被指数(Enhanced vegetation index, EVI)[28]、比值植被指数(Ratio vegetation index, RVI)[29]、土壤调节植被指数(Soil-adjusted vegetation index, SAVI)[30]和修正土壤调节植被指数(Modified soil-adjusted vegetation index, MSAVI)[31]。每个年份最终生成包含 12 个月度时间步、10 个特征通道(4 个反射率波段 B2、B3、B4、B8 和 6 个指数)的光学时序张量,总计 120 维输入数据。

1.2.2 Sentinel-1 SAR 数据

Sentinel-1 SAR 数据的预处理从轨道选择开始,本研究仅采用升轨数据。随后进行基于物理模型的辐射地形校正,利用 30 m 分辨率的 SRTM GL1 数字高程模型,将原始 Sigma0 后向散射系数转换为 Gamma0[32]。在此基础上,应用 Refined Lee 滤波抑制斑点噪声[33]。

基于校正和滤波后的同极化后向散射系数(Vertical-vertical polarization backscatter coefficient, VV)与交叉极化后向散射系数(Vertical-horizontal polarization backscatter coefficient, VH),进一步计算极化差(VV minus VH, VV-VH)和雷达植被指数(Radar vegetation index, RVI)作为增强特征。在时序构建方面,将全年划分为 24 个 15 d 间隔的时间

步，对每个时间步，取中心时间前后各 15 d 内的所有有效影像的中位数进行合成。每个年份最终生成包含 24 个时间步、4 个特征通道（VV、VH、VV－VH、RVI）的雷达时序张量，共计 96 维输入数据。主要参数计算方法为

$$V_V - V_H = V_{V-H} \tag{1}$$

$$\sigma_{VV} = 10^{\frac{V_V}{10}} \tag{2}$$

$$\sigma_{VH} = 10^{\frac{V_H}{10}} \tag{3}$$

$$\mathrm{RVI} = \frac{4\sigma_{VH}}{\sigma_{VV} + \sigma_{VH}} \tag{4}$$

式中　V_V——基于辐射地形校正与 Refined Lee 滤波后的同极化后向散射系数，dB

V_H——基于辐射地形校正与 Refined Lee 滤波后的交叉极化后向散射系数，dB

V_{V-H}——极化差

σ_{VV}——VV 由分贝换算得到的线性功率

σ_{VH}——VH 由分贝换算得到的线性功率

RVI——雷达植被指数

由于 RVI 需在线性功率域计算，故先按式（2）、（3）将 dB 形式的 V_V、V_H 转换为线性功率σ_{VV}、σ_{VH}，再代入式（4）。

1.3　样本采集

本研究构建了涵盖甘蔗、水稻、玉米、其他耕地、林地、灌木、草地、不透水面及水体在内的 9 类地物分类体系。采用分层随机抽样与多源数据交叉核验相结合的策略，逐年独立构建了 2019—2025 年的样本数据集，以减少时序自相关偏差，并保持样地属性与影像时相的一致性。样本来源主要包括 3 部分：广西甘蔗野外调查数据及甘蔗参考数据集[34－35]；CLCD 土地覆盖数据集[36]、中国水稻分布数据集[37]及中国玉米物候数据集[38]；基于 Google Earth 高分辨率影像与 Sentinel－2 L2A 月度合成影像的目视判读样本。在样本构建过程中，以历年外业实测点为空间锚点，对照当年影像逐年核查样点状态，剔除地类变更或时相不符的异常点。2019—2025 年间，每年采集甘蔗训练样本 1 600～2 000 个，8 类非甘蔗训练样本每类各 500 个，验证样本独立于训练样本采集（表 1）。以 2019 年为例，样本点的空间分布如图 2 所示。

基于上述样本，研究提取了 2019 年 9 类地物在月度尺度上的全部光学与雷达特征时间序列（图 3）。其中光学特征包括 4 个反射率波段（B2、B3、B4、B8）和 6 个遥感指数（NDVI、NDWI、EVI、RVI、SAVI、MSAVI），雷达特征包括 VV、VH、VV－VH 和雷达植被指数（RVI）。

表 1　崇左市 2019—2025 年样本数量汇总

Tab. 1　Sample sizes for Chongzuo City from 2019 to 2025

年份	训练样本数		验证样本数	
	甘蔗	非甘蔗	甘蔗	非甘蔗
2019	1 600	4 000	651	1 481
2020	1 600	4 000	667	1 565
2021	1 900	4 000	728	1 333
2022	1 900	4 000	711	1 395
2023	1 882	4 000	702	1 333
2024	1 942	4 000	698	1 420
2025	1 898	4 000	742	1 391

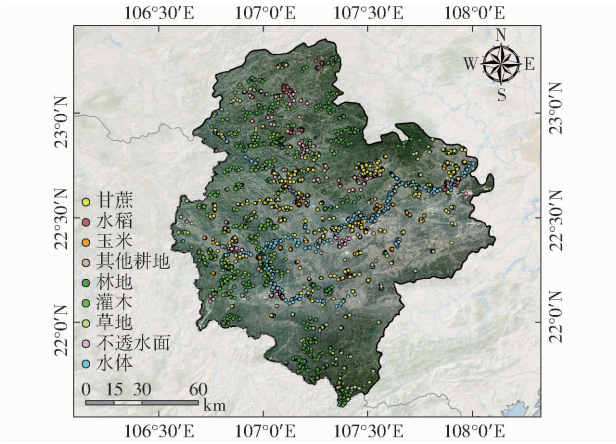

图 2　样本点分布（2019 年）

Fig. 2　Distribution of sample points (2019)

从光学波段（图 3a～3d）来看，甘蔗在可见光波段（B2、B3、B4）的反射率在伸长期（7—9 月）出现明显下降，与其他作物形成区分；近红外波段（B8）的反射率则在伸长期达到峰值（0.35～0.40），反映了甘蔗冠层叶面积指数的快速增长。从植被指数（图 3e、3g、3i、3j）来看，NDVI、EVI、SAVI、MSAVI 均呈现出较为一致的物候节律：1—4 月因收获与播种处于低值，5 月分蘖期后快速上升，8—9 月伸长期达到峰值，随后在成熟期缓慢回落，其中 SAVI 与 MSAVI 同属土壤调节类指数，在以植被覆盖为主的场景下两者曲线形态接近，本研究保留二者以保持光学特征集的完整性，其细微差异主要体现在低植被覆盖地物上；NDWI（图 3f）整体为负值并逐月下降，反映了甘蔗冠层含水率的相对变化；光学 RVI（图 3h）的变化趋势与 NDVI 接近。

从雷达特征（图 3k～3n）来看，VH 通道的物候节律最为清晰：4 月降至最低（约 －18.5 dB），6—7 月上升至峰值（约 －14.8 dB），与 NDVI 变化基本同步；VV 的年内变化幅度小于 VH；VV－VH 在分蘖至伸长期相对稳定，反映了甘蔗冠层结构的连续性；雷达植被指数（RVI）在伸长期至成熟过程中表现为缓慢上升，对甘蔗与背景地物有一定的区分作用。

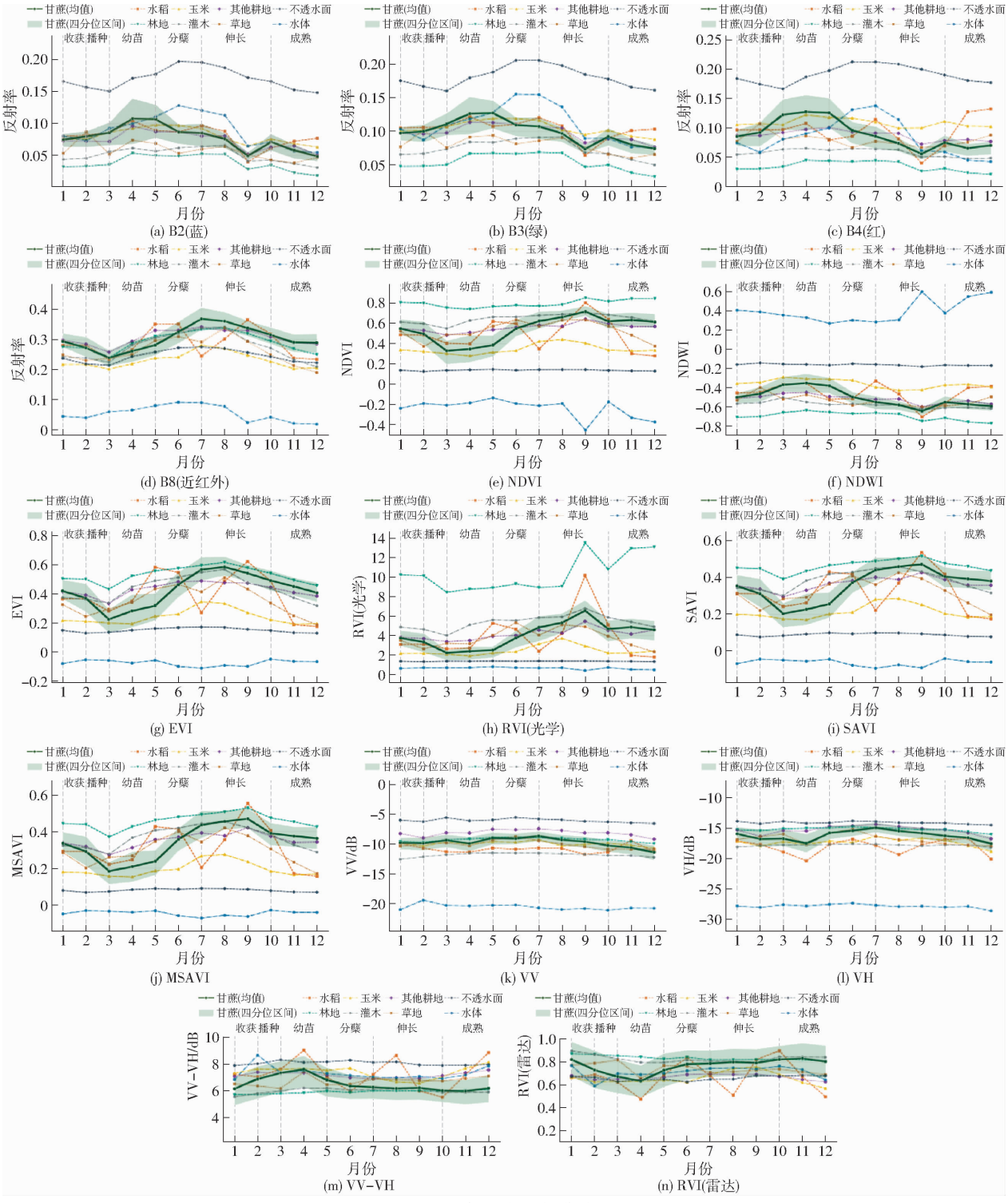

图 3 各地类物候曲线

Fig. 3 Phenological curves of different land cover types

总体来看,甘蔗与水稻、玉米等物候相似作物及林地、灌木等背景地物在多数光学与雷达特征上均呈现可区分的差异,表明 Sentinel 多源时序数据为 9 类地物的判别提供了较为充分的特征基础。

2 研究方法

2.1 技术路线

本文的研究路线包括 3 个主要环节(图 4):基于 Google Earth Engine 平台对 Sentinel 数据进行获取及预处理,分别得到 Sentinel-2 光学时序特征和 Sentinel-1 雷达时序特征,同时基于外业数据、Google Earth 影像以及相关土地覆盖数据集制备实验样本;基于 LSTM 的分类架构,进行数据源对比和历史信息利用策略对比实验;采用独立验证样本对各方案的分类结果进行精度评估。

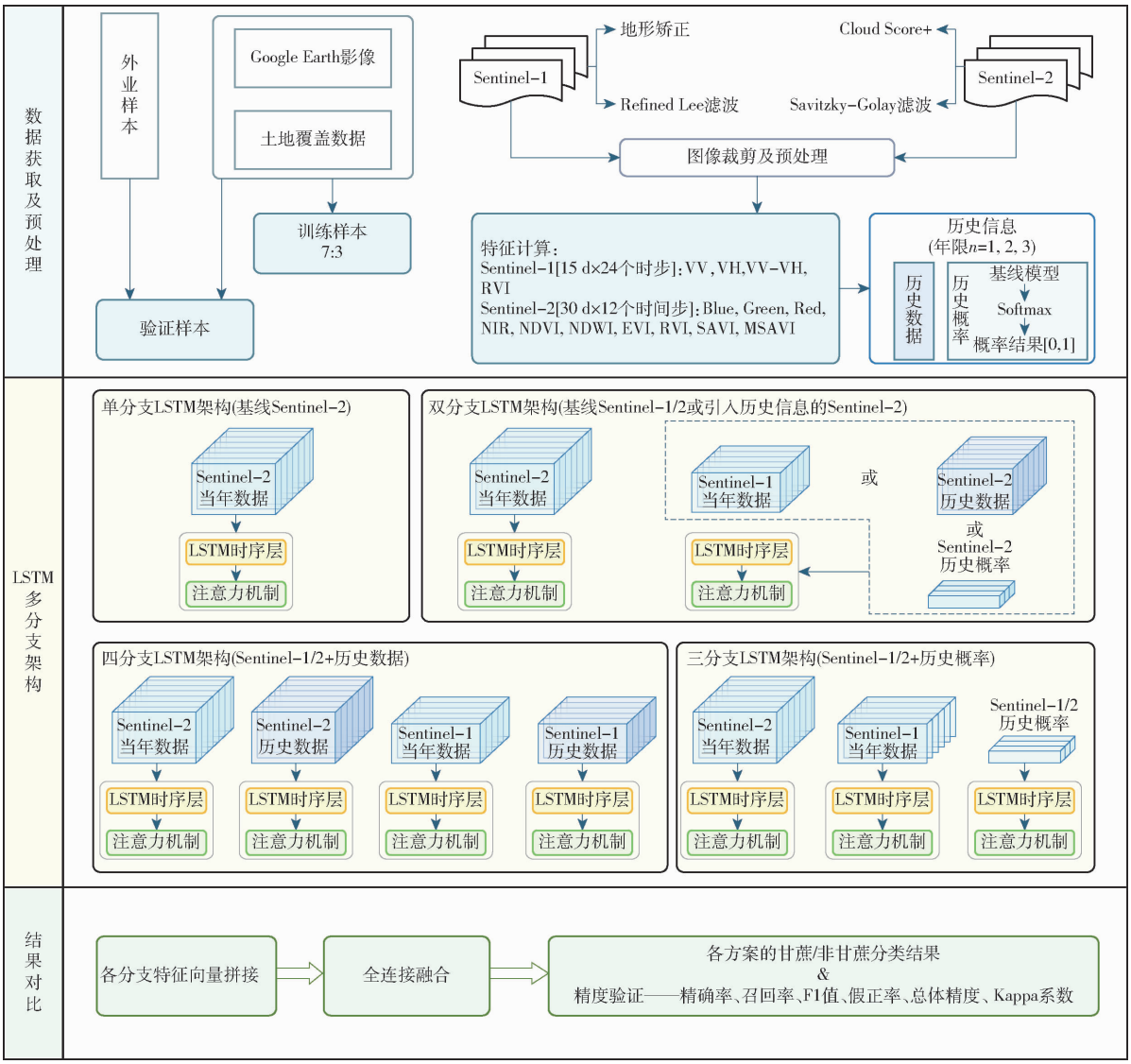

图 4　技术路线图

Fig. 4　Technical flowchart

2.2　分类模型架构

本研究根据数据源特性与历史信息引入方式，设计了 3 种 LSTM 架构变体(图 4)。各架构的基础处理模块均由 LSTM 时序层、注意力机制和全连接融合层组成。其中，单分支 LSTM 架构作为基线模型，用于仅依赖 Sentinel－2 光学数据的分类任务，通过 LSTM 层提取 12 个月度时间步的多维特征，并利用注意力机制进行动态加权。针对多源数据或历史信息的引入，研究进一步设计了双分支 LSTM 架构：一类用于 Sentinel－1 与 Sentinel－2 的基线融合分类，由两条独立分支分别处理光学(12 步 × 10 维)与雷达(24 步 ×4 维)时序，再将特征向量拼接融合；另一类用于 Sentinel－2 引入历史信息的场景，由两条分支分别建模当年时序与前 n 年的历史先验序列。

当 Sentinel－1/2 融合数据进一步引入历史信息时，研究构建了多分支 LSTM 架构。其中，历史数据方案采用四分支结构，分别对当年和历史的光学、雷达数据进行独立建模；历史概率方案采用三分支结构，包含当年光学、当前雷达和历史概率序列 3 条分支。采用独立分支的目的是分别提取不同模态和不同年份的时序特征，避免特征直接堆叠造成的时相语义混淆。其中，每一年份的归属概率基于该年份数据独立生成，取值范围为[0，1]，表示像元属于甘蔗的可能性。在多年回溯($n>1$)场景下，各年概率按时间顺序拼接为短序列输入。

为提升分类器对相似作物的区分能力，各模型均采用 9 类地物的多分类训练，通过明确的类别边界限度降低甘蔗与水稻、玉米等物候相似作物间的混淆。模型训练采用 AdamW 优化器，其在 Adam 的基础上将权重衰减与梯度更新过程解耦，直接以独立项的形式作用于模型参数，避免了 Adam 中自适应学习率与 L2 正则相互耦合而导致的正则化效果减弱问题，使权重衰减能够稳定地发挥抑制过拟合

的作用[39]。同时,AdamW 沿用了 Adam 的一阶矩与二阶矩自适应估计机制,可根据各参数的梯度统计自动调节更新步长,对处理多源遥感时序数据中量纲差异较大、稀疏且非平稳的特征较为适用。本研究中,AdamW 的初始学习率设为 1×10^{-3},权重衰减系数设为 1×10^{-3}。

为进一步稳定训练并提升泛化能力,研究采用了 3 项配套策略:一是余弦退火学习率调度,使学习率沿余弦曲线平滑衰减,训练前期以较大步长快速收敛,后期以较小步长精细优化;二是梯度裁剪,将每次反向传播的梯度全局范数限制在 1.0 以内,抑制 LSTM 训练中常见的梯度爆炸现象;三是标签平滑交叉熵损失(平滑系数 0.1),缓解模型对训练标签的过度自信,降低对噪声标签和混合像元的敏感性。此外,训练过程中还引入了高斯噪声注入与随机通道丢弃作为在线数据增强手段,进一步提升模型对观测扰动的稳健性。

在超参数配置方面,LSTM 网络采用单层结构,隐藏层维度在 48 ~ 64 区间内根据验证集性能自动调优,训练轮次上限设为 50 轮并引入早停机制,即验证损失连续 10 轮不下降时停止训练,批次大小固定为 32。3 种 LSTM 架构变体的差异仅在于输入特征的组合方式,而 LSTM + Attention 的基础结构、AdamW + 余弦退火 + 梯度裁剪 + 标签平滑的训练策略以及超参数设置在所有变体中保持一致,以保证不同方案精度对比的公平性。各模型的训练样本按 7∶3随机划分为训练集和内部验证集,分别用于模型训练与超参数调优;最终分类精度采用与训练样本相互独立的验证样本进行评估。

2.3 实验设计

本研究设计了 14 组对比实验(表 2),由于 $n=3$ 的历史信息方案需要前 3 年的基线分类结果作为输入,且研究区 Sentinel 数据质量可用年份自 2019 年起始,因此所有方案的精度评估统一在 2022—2025 年 4 个目标年份上进行,以保证不同方案对比的公平性。基线方案和 $n=1$ 方案在技术上可覆盖更多年份,但仅 2022—2025 年能满足 $n=3$ 方案的历史回溯需求,故以该时间窗口作为所有实验的统一评估范围。

实验设计遵循控制变量法原则:E1 与 E2 构成两种基线对比;E3 ~ E5 和 E6 ~ E8 在固定数据源的条件下考察历史数据回溯年限($n=1,2,3$)的影响;E9 ~ E11 和 E12 ~ E14 对应考察历史概率回溯年限的影响;相同数据源、相同回溯年限下的历史数据方案与历史概率方案对比(如 E5 与 E11、E8 与 E14),评估两种引入方式的相对优劣。

表 2 实验设计方案

Tab. 2 Experimental design

方案	数据源	历史信息(n 为年限)
E1	Sentinel - 2	—
E2	Sentinel - 1/2	—
E3	Sentinel - 2	历史数据($n=1$)
E4	Sentinel - 2	历史数据($n=2$)
E5	Sentinel - 2	历史数据($n=3$)
E6	Sentinel - 1/2	历史数据($n=1$)
E7	Sentinel - 1/2	历史数据($n=2$)
E8	Sentinel - 1/2	历史数据($n=3$)
E9	Sentinel - 2	历史概率($n=1$)
E10	Sentinel - 2	历史概率($n=2$)
E11	Sentinel - 2	历史概率($n=3$)
E12	Sentinel - 1/2	历史概率($n=1$)
E13	Sentinel - 1/2	历史概率($n=2$)
E14	Sentinel - 1/2	历史概率($n=3$)

注:"—"表示无历史信息加入。

2.4 精度评价

对各方案的分类结果采用 6 项指标进行评估:精确率、召回率、F1 值、假正率、总体精度和 Kappa 系数。精度评估采用独立于训练集的验证样本,对每个实验的每个年份独立计算上述指标。

3 结果与分析

3.1 单年基线方案结果

表 3 展示了 14 组方案在 2022—2025 年间的 4 年平均精度。结果显示,仅使用 Sentinel - 2 光学数据的基线方案(E1)F1 值为 83.1%,Kappa 系数为 0.726;融合 Sentinel - 1 与 Sentinel - 2 的基线方案(E2)F1 值为 82.3%,Kappa 系数为 0.715。可见在

表 3 各方案 2022—2025 年平均精度

Tab. 3 Average accuracy of each solution from 2022 to 2025

方案	精确率/%	召回率/%	F1 值/%	假正率/%	总体精度/%	Kappa 系数
E1	73.7	95.3	83.1	17.5	86.8	0.726
E2	75.5	91.9	82.3	16.4	86.3	0.715
E3	74.9	96.1	84.1	16.7	87.7	0.743
E4	81.6	95.7	88.0	11.4	91.0	0.809
E5	83.7	96.9	89.8	9.7	92.5	0.839
E6	80.5	94.9	87.1	11.8	90.4	0.795
E7	83.1	95.0	88.6	10.0	91.7	0.821
E8	86.2	95.6	90.6	7.9	93.3	0.854
E9	76.6	95.3	84.9	15.0	88.5	0.758
E10	78.4	96.0	86.3	13.6	89.7	0.781
E11	79.6	96.1	87.0	12.8	90.2	0.793
E12	78.0	96.0	86.0	14.0	89.4	0.776
E13	77.9	97.2	86.3	14.5	89.5	0.780
E14	85.2	95.6	90.1	8.6	92.8	0.845

单年基线场景下，引入雷达数据并未带来精度增益，反而略低于纯光学方案。原因在于，Sentinel－1 SAR 对冠层密度、茎秆排列等结构特征较为敏感，但对光谱物候特征的表征能力有限；在单年场景下，其判别能力还受到斑点噪声及对作物类型区分度较低的影响。在样本量有限的情况下，将 SAR 与光学特征在空间维度上简单堆叠，难以在单年分类中产生协同增益。

3.2　历史信息引入方式与回溯年限的影响

表 3 显示，引入历史信息后 12 组方案（E3～E14）在 F1 值与 Kappa 系数上均优于同源基线方案（E1 与 E2）。该结果充分表明，无论采取历史数据还是历史概率均能提升模型对 Sentinel 时序数据的利用程度，进而提高甘蔗识别精度。

从回溯年限来看，历史数据方案精度随 n 的增加而递增：仅使用 Sentinel－2 数据的方案在 n 由 0 增至 3 时，F1 值由 83.1% 稳步提升至 89.8%（增长 6.7 个百分点）；Sentinel－1/2 融合方案中的 F1 值由 82.3% 升至 90.6%（增长 8.3 个百分点）。历史概率方案在 $n=1$～3 区间也呈现相似的增长趋势。两种引入方式均在 $n=3$ 时取得最高精度。这一结果表明，3 年的回溯长度可完整覆盖广西甘蔗“一年新植、两年宿根”的农艺周期，为模型提供较为完整的时间先验，也可作为其他多年生作物遥感监测的回溯年限参考。

对比相同条件下的两种引入方式，历史数据方案在整体精度上略优于历史概率方案。在仅利用光学数据且 $n=3$ 时，历史数据方案的 F1 值（89.8%）比历史概率方案（87.0%）高约 2.8 个百分点；在 Sentinel－1/2 融合场景下，两者差距缩小至 0.5 个百分点。这表明随着输入特征维度的增加和模型复杂度的提升，两种方式的提取效能趋于等效。

此外，研究发现，SAR 数据在不同场景下的贡献呈现出明显的反转。单年基线场景下，Sentinel－1/2 融合方案（E2）略低于单光学方案（E1）；引入 3 年历史信息后，融合方案（E8、E14）的精度均明显高于对应的光学方案（E5、E11）。这一现象反映出 SAR 数据对多年生作物冠层结构延续性较为敏感的特点：单年场景下缺乏可比较的时间基准，SAR 的优势难以体现；引入 3 年历史信息后，SAR 可作为结构延续性的判断依据，与光学的物候信息形成互补。

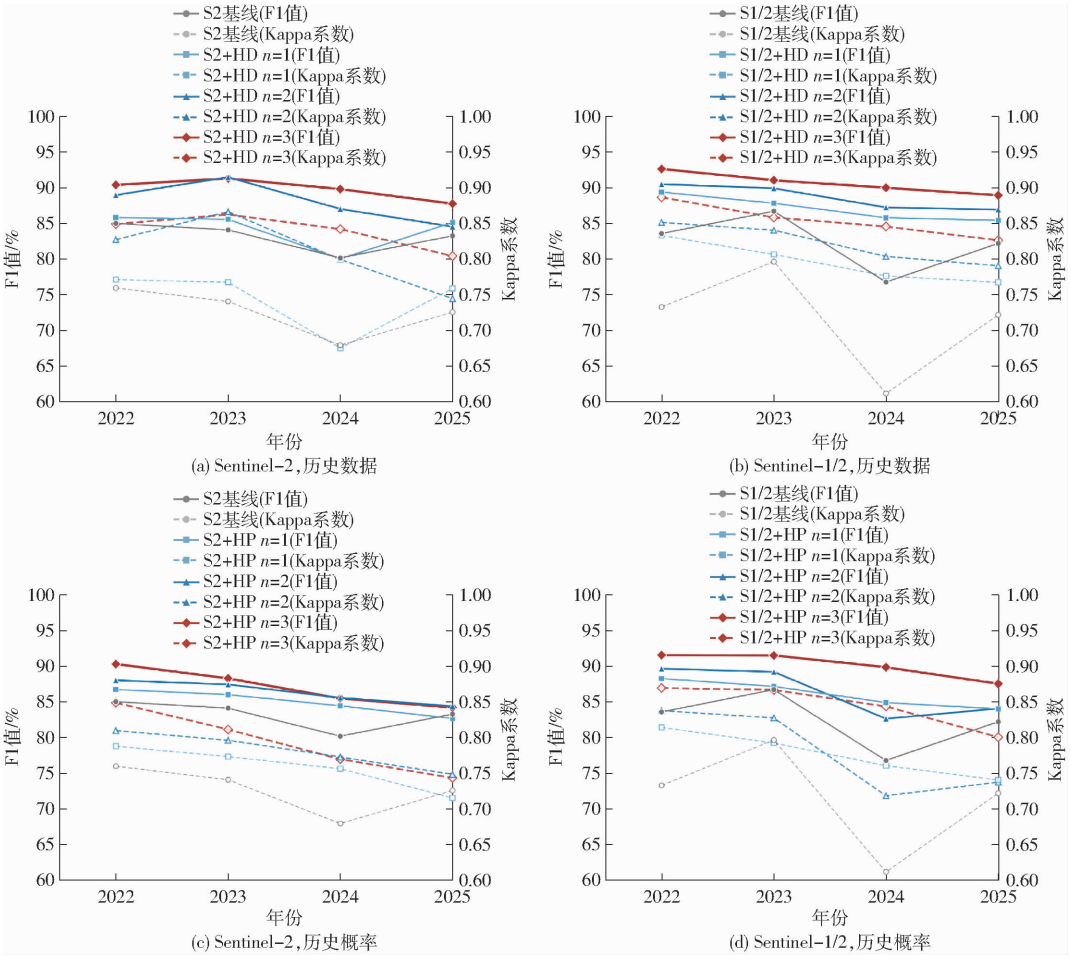

图 5　基于不同数据源的历史信息引入方式对比结果

Fig. 5　Comparison results of historical information integration methods based on different data sources

1/2,HD 表示历史数据,HP 表示历史概率)反映了 14 组方案在 2022—2025 年的精度波动情况。结果显示,基线方案的年际波动较大,E2 方案在 2024 年的 F1 值降至 76.7%,说明 Sentinel-1/2 基线在特定气象条件下精度不稳定。相比之下,引入 3 年历史信息后的最优方案(E8)在 4 个年份中的 F1 值稳定在 88.9% ~92.6% 之间,年际波动明显减弱。这表明甘蔗种植的多年连续性不仅可以提升精度的上限,也能在一定程度上稳定模型表现:当目标年份的光学观测受云覆盖或极端天气影响时,历史先验能够提供补充信息,降低分类结果对单一年份观测质量的依赖。

图 5 结果还显示,各方案的 Kappa 系数在 2022—2025 年间整体呈下降趋势。基线方案下降较为明显,如 E2 方案的 Kappa 系数由 2022 年的 0.732 降至 2024 年的 0.611;引入 3 年历史信息的 E8 方案下降幅度较小,由 2022 年的 0.886 降至 2025 年的 0.826。造成这一趋势的原因可能有以下几方面:一是研究区近年来种植结构调整加快,甘蔗与其他经济作物之间的轮作和转换增多,类别分布的变化对 Kappa 系数有一定影响;二是部分年份的气象条件不利于 Sentinel-2 的稳定观测,单年方案受影响较为明显;三是验证样本的类别构成在不同年份存在差异,也会对 Kappa 系数的统计结果产生影响。但从对比来看,引入 3 年历史信息后,Kappa 系数的年际下降幅度由基线方案的 0.10 ~0.13 减小至 0.05 左右,说明历史先验信息有助于提升模型的年际稳定性。

3.3 崇左市 2022—2025 年甘蔗分布

基于最优方案 E8,本研究生成了崇左市 2022—2025 年 10 m 分辨率甘蔗空间分布图(图 6)。结果显示,甘蔗集中分布在江州区、扶绥县、龙州县以及大新县东部和宁明县北部,空间格局与当地农业发展规划基本一致。4 年观测期内,核心连片种植区的空间稳定性较高,而边缘地块则呈现一定比例的种植切换。为更直观地展示局部地块的年际动态,图 6 中进一步选取 s1 和 s2 两个典型区域进行放大对比:s1 区域(位于江州区)在 2022—2025 年间甘蔗种植范围逐年扩张,反映了种植规模的持续增长;s2 区域(位于扶绥县)则呈现退种或轮作切换的趋势,对应种植规模逐步减小。两个区域均在 4 个年份中保持相同的空间范围,可清晰反映方法在捕捉地块尺度时空动态变化方面的能力。这一时空特征与广西“一年新植、两年宿根”的 3 年农艺周期相符,即同一地块在完整周期内持续种植甘蔗,周期结束后进行翻蔸重种或开展作物轮作。

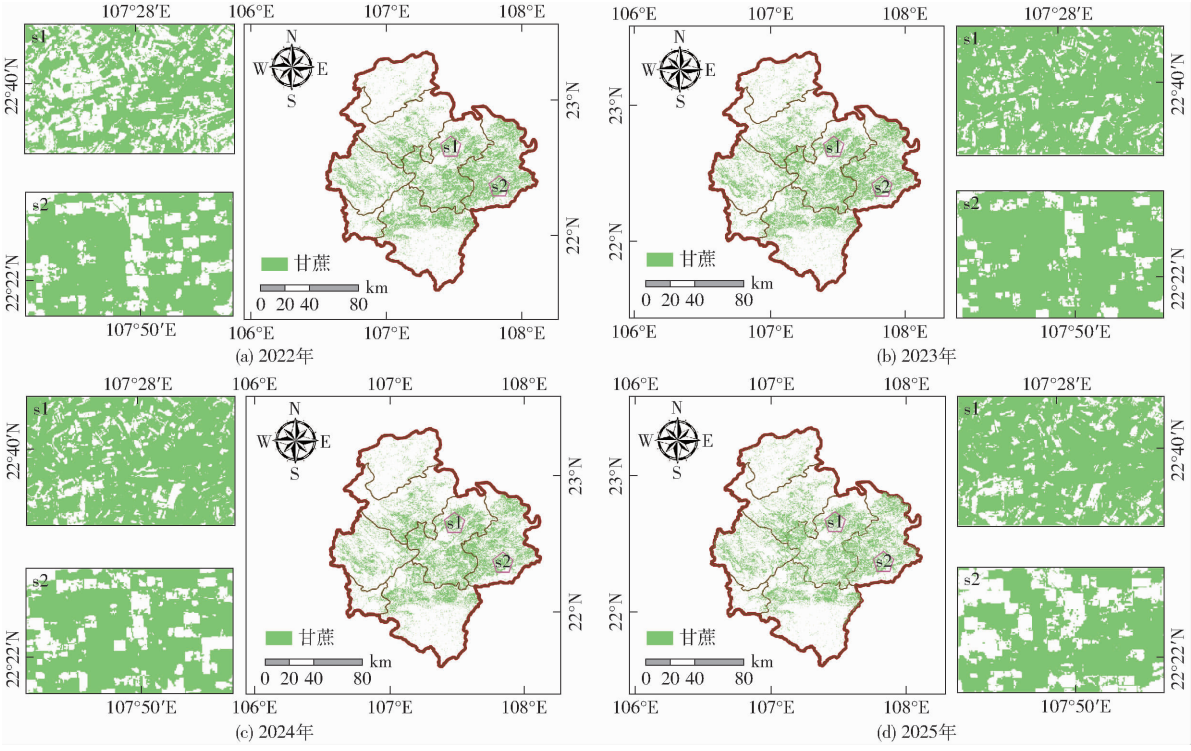

图 6 2022—2025 年崇左市甘蔗分布结果

Fig. 6 Sugarcane distribution results in Chongzuo City from 2022 to 2025

4 讨论

本研究的 14 组对比实验揭示了 Sentinel-1 对光学数据的贡献具有明显的条件性特征,历史数据与历史概率两种引入方式的精度差异随融合复杂度而收敛,最佳回溯年限与甘蔗农艺周期相一致。以

下围绕这 3 个方面展开探讨。

(1)SAR 数据贡献的条件性机理

Sentinel－1 在基线与历史信息场景下呈现出明显不同的贡献模式。基线场景下,Sentinel－1/2 融合方案 F1 值略低于 Sentinel－2 方案,但在引入 3 年历史信息后,融合方案的精度明显高于仅光学方案。这一现象与 SAR 数据的信息特性有关。Sentinel－1 对植被结构敏感,但对光谱物候变化不敏感,加之受斑点噪声影响,单年条件下对作物类型的区分能力有限。但是,SAR 对多年生作物冠层结构的延续性较为敏感:若同一地块在相邻年份持续种植甘蔗,其 SAR 时序往往呈现相近的季节性变化;反之,若发生作物轮换,SAR 时序模式将产生显著差异。这一特征需要与历史先验信息配合才能发挥作用,单年分类下的 SAR 缺乏可对比的时间基准,其价值难以显现;引入前 3 年的历史信息后,SAR 可用于判断当前年份的结构是否延续了历史模式,从而与光学数据的物候信息形成互补。该机理对其他冠层结构相对稳定的多年生作物的 SAR－光学融合分类具有一定参考意义。

(2)不同历史先验引入策略的效能权衡

历史数据方案的精度在 $n=3$ 条件下略高于历史概率方案,但两者的差距随数据源融合复杂度的增加而缩小(仅 Sentinel－2 场景 F1 值差距为 2.8 个百分点,Sentinel－1/2 融合场景仅为 0.5 个百分点)。从信息特征来看,两种历史信息引入方式都带有一定噪声,但噪声的性质不同。历史数据的噪声主要来自遥感观测过程,包括云掩膜残留、大气校正偏差、SAR 斑点噪声以及极端气象导致的物候异常等。这类噪声较为随机,年际间相对独立,模型在处理多年数据时可在一定程度上相互抵消。历史概率的噪声则来源于基线模型的判别误差。本研究中,Sentinel－2 与 Sentinel－1/2 基线模型的 F1 值分别为 83.1% 和 82.3%,基线模型对某些地物的误判在不同年份间往往具有一致性,使得历史概率中的噪声带有系统性偏差,多年叠加后难以通过平均的方式消除,甚至可能被下游模型当作稳定的先验加以利用。此外,历史数据保留了完整的多维时序信息,模型可在原始特征层面直接进行跨年对比;而历史概率将多维信息压缩为单一标量,会丢失部分有助于跨年判别的细节。当数据源融合复杂度增加时,模型需同时处理光学与雷达的跨模态融合,学习负担加大,历史数据引入的额外高维信息优势会被部分抵消,因此两种方式的精度趋于接近。尽管历史数据在精度上略有优势,但历史概率方案的实用价值仍不可忽视:其历史分支仅需处理 1 维概率序列而非完整的多维时序张量,在计算量、存储成本和推理效率方面均明显优于历史数据方案,在大规模业务化应用场景中具有重要价值。

(3)回溯年限与农艺周期的耦合关系

回溯年限 $n=3$ 在两种引入方式和两种数据源下均取得最优精度,这与广西“一年新植、两年宿根”的典型 3 年甘蔗种植周期较为一致。3 年的回溯长度恰好能够覆盖一个完整的种植周期,既不因过短而只反映周期的一部分,也不因过长而引入周期外的无关信息。这一规律也可为其他多年生作物的回溯年限设定提供参考,即以作物主要种植周期的长度作为最佳 n 的参考值。

5　结论

(1)本研究面向多年生作物时序遥感分类中历史先验信息利用不足的问题,基于 Sentinel－2 L2A 与 Sentinel－1 SAR 时序数据,构建了融合多年历史信息的 LSTM 多分支分类框架,实现了广西崇左市 2022—2025 年甘蔗种植区的连续、稳定提取,验证了该框架在多年尺度甘蔗制图中的可行性,其方法思路可推广至其他多年生作物的时序遥感分类任务。

(2)引入多年历史信息能够显著提升 Sentinel 时序数据的甘蔗分类精度并改善年际稳定性,最优方案(Sentinel－1/2 融合结合 3 年历史数据)2022—2025 年的 4 年平均 F1 值达到 90.6%、Kappa 系数为 0.854,较同源单年基线 F1 值提升约 8.3 个百分点;其中 Sentinel－1 数据的贡献具有条件性,基线场景下融合 SAR 未能带来精度提升,但在 3 年历史信息支持下,SAR 的结构敏感性与时间先验形成互补,使融合方案明显优于仅光学方案。

(3)历史数据与历史概率两种引入方式在 3 年回溯条件下均可取得 F1 值不低于 90% 的精度,前者在精度上略占优势(差距 0.5～2.8 个百分点),后者在计算与存储效率上更具优势,两者在不同应用场景中各具适用性。

(4)回溯年限与分类精度呈单调递增关系,$n=3$ 时取得最优,该最佳年限与广西“一年新植、两年宿根”的典型 3 年甘蔗种植周期相吻合,揭示了模型有效历史信息长度与作物农艺周期之间的对应关系。

参　考　文　献

[1]　Verma K K, Song X P, Yadav G, et al. Impact of agroclimatic variables on proteogenomics in sugar cane (*Saccharum spp.*)

plant productivity[J]. ACS Omega, 2022, 7(27): 22997 – 23008.
[2] FAO. World food and agriculture: statistical yearbook 2020[M]. Rome: FAO, 2020.
[3] 詹鹏举. 全球背景下的广西糖业发展现状及对策[D]. 南宁: 广西大学, 2013.
Zhan Pengju. Development status and countermeasures of Guangxi's sugar industry under the global background[D]. Nanning: Guangxi University, 2013. (in Chinese)
[4] Chen W, Shi Y, Yan W, et al. Large-scale long-term accurate mapping of sugarcane fields using a combination of active and passive remote sensing data[J]. International Journal of Remote Sensing, 2025, 46(21): 8321 – 8345.
[5] Yin D, Yan J, Li F, et al. Evaluation of global gridded crop models in simulating sugarcane yield in China[J]. Atmospheric and Oceanic Science Letters, 2023, 16(2): 100329.
[6] Joshi A, Pradhan B, Gite S, et al. Remote-sensing data and deep-learning techniques in crop mapping and yield prediction: a systematic review[J]. Remote Sensing, 2023, 15(8): 2014.
[7] Liu X, Zhai H, Shen Y, et al. Large-scale crop mapping from multisource remote sensing images in Google Earth Engine[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 414 – 427.
[8] Wu B, Zhang M, Zeng H, et al. Challenges and opportunities in remote sensing-based crop monitoring: a review[J]. National Science Review, 2023, 10(4): nwac290.
[9] Baranoski G V G. Foliar spectral responses of sugarcane and maize: how comparable are they? [C]// Remote Sensing for Agriculture, Ecosystems, and Hydrology XXV: Vol. 12727. Bellingham: SPIE, 2023: 127270V.
[10] Som-Ard J, Atzberger C, Izquierdo-Verdiguier E, et al. Remote sensing applications in sugarcane cultivation: a review[J]. Remote Sensing, 2021, 13(20): 4040.
[11] Xu F, Wang Z, Lu G, et al. Sugarcane ratooning ability: research status, shortcomings, and prospects[J]. Biology, 2021, 10(10): 1052.
[12] Russwurm M, Körner M. Temporal vegetation modelling using long short-term memory networks for crop identification from medium-resolution multi-spectral satellite images[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2017: 1496 – 1504.
[13] Meroni M, D'Andrimont R, Vrieling A, et al. Comparing land surface phenology of major European crops as derived from SAR and multispectral data of Sentinel – 1 and – 2[J]. Remote Sensing of Environment, 2021, 253: 112232.
[14] 王丽美, 姜永涛, 张彩丽, 等. 基于 Sentinel – 1/2 数据优势特征选择的冬小麦层级提取方法[J]. 农业机械学报, 2026, 57(7): 295 – 307.
Wang Limei, Jiang Yongtao, Zhang Caili, et al. Winter wheat hierarchical extraction method based on dominant feature selection from Sentinel – 1/2 data[J]. Transactions of the Chinese Society for Agricultural Machinery, 2026, 57(7): 295 – 307. (in Chinese)
[15] Di Tommaso S, Wang S, Strey R, et al. Mapping sugarcane globally at 10 m resolution using Global Ecosystem Dynamics Investigation (GEDI) and Sentinel – 2[J]. Earth System Science Data, 2024, 16(10): 4931 – 4947.
[16] Maleki S, Baghdadi N, Bazzi H, et al. Machine learning-based summer crops mapping using Sentinel – 1 and Sentinel – 2 images[J]. Remote Sensing, 2024, 16(23): 4548.
[17] Kou W, Shen Z, Liu D, et al. Crop classification methods and influencing factors of reusing historical samples based on 2D – CNN[J]. International Journal of Remote Sensing, 2023, 44(10): 3278 – 3305.
[18] Lin C, Zhong L, Song X, et al. Early- and in-season crop type mapping without current-year ground truth: generating labels from historical information via a topology-based approach[J]. Remote Sensing of Environment, 2022, 274: 112994.
[19] Wang H, Ye Z, Wang Y, et al. Improving the crop classification performance by unlabeled remote sensing data[J]. Expert Systems with Applications, 2024, 236: 121283.
[20] Wang Q, Yang B, Li L, et al. Within-season crop identification by the fusion of spectral time-series data and historical crop planting data[J]. Remote Sensing, 2023, 15(20): 5043.
[21] Račič M, Oštir K, Zupanc A, et al. Multi-year time series transfer learning: application of early crop classification[J]. Remote Sensing, 2024, 16(2): 270.
[22] 谢鑫昌, 杨云川, 田忆, 等. 基于遥感的广西甘蔗种植面积提取及长势监测[J]. 中国生态农业学报(中英文), 2021, 29(2): 410 – 422.
Xie Xinchang, Yang Yunchuan, Tian Yi, et al. Remote sensing based sugarcane planting area extraction and growth monitoring in Guangxi[J]. Chinese Journal of Eco-Agriculture, 2021, 29(2): 410 – 422. (in Chinese)
[23] Hong Y, Xie T, Luo L, et al. Area extraction and growth monitoring of sugarcane from multi-source remote sensing images under a polarimetric SAR data compensation based on buildings[J]. Geo-Spatial Information Science, 2025, 28(3): 831 – 848.
[24] Pasquarella V J, Brown C F, Czerwinski W, et al. Comprehensive quality assessment of optical satellite imagery using weakly supervised video learning[C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2023: 2125 – 2135.

(下转第 103 页)

21768.
[31] Periasamy S. Significance of dual polarimetric synthetic aperture radar in biomass retrieval: an attempt on Sentinel－1[J]. Remote Sensing of Environment, 2018, 217: 537－549.
[32] Nasirzadehdizaji R, Balik S F, Abdikan S, et al. Sensitivity analysis of multi-temporal Sentinel－1 SAR parameters to crop height and canopy coverage[J]. Applied Sciences, 2019, 9(4): 655.
[33] Yamada Y. Preliminary study on the radar vegetation index (RVI) application to actual paddy fields by ALOS/PALSAR full-polarimetry SAR data[J]. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2015, 40: 129－131.
[34] Kim Y, Jackson T, Bindlish R, et al. Retrieval of wheat growth parameters with radar vegetation indices[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 11(4): 808－812.
[35] Mandal D, Ratha D, Bhattacharya A, et al. A radar vegetation index for crop monitoring using compact polarimetric SAR data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(9): 6321－6335.
[36] Schlund M, Erasmi S. Sentinel－1 time series data for monitoring the phenology of winter wheat[J]. Remote Sensing of Environment, 2020, 246: 111814.
[37] 代天金，陈俊英，郭佳奇，等. 不同植被覆盖条件下 Sentinel－1/2 数据融合监测土壤含盐量模型研究[J]. 农业机械学报，2025，56(8)：32－41.
Dai Tianjin, Chen Junying, Guo Jiaqi, et al. Monitoring of soil salt content during different growth periods of crops based on Sentinel－1/2[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(8): 32－41. (in Chinese)
[38] Bogdanovski O P, Svenningsson C, Månsson S, et al. Yield prediction for winter wheat with machine learning models using Sentinel－1, topography, and weather data[J]. Agriculture, 2023, 13(4): 813.
[39] Kumar P, Prasad R, Gupta D, et al. Estimation of winter wheat crop growth parameters using time series Sentinel－1A SAR data[J]. Geocarto International, 2018, 33(9): 942－956.

(上接第 64 页)

[25] Savitzky A, Golay M J E. Smoothing and differentiation of data by simplified least squares procedures[J]. Analytical Chemistry, 1964, 36(8): 1627－1639.
[26] Rouse J W, Haas R H, Schell J A, et al. Monitoring vegetation systems in the Great Plains with ERTS[C]// Proceedings of the Third Earth Resources Technology Satellite－1 Symposium: NASA SP－351. Washington: NASA, 1974: 309－317.
[27] Gao B C. NDWI: a normalized difference water index for remote sensing of vegetation liquid water from space[J]. Remote Sensing of Environment, 1996, 58(3): 257－266.
[28] Huete A, Didan K, Miura T, et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices[J]. Remote Sensing of Environment, 2002, 83(1－2): 195－213.
[29] Jordan C F. Derivation of leaf-area index from quality of light on the forest floor[J]. Ecology, 1969, 50(4): 663－666.
[30] Huete A R. A soil-adjusted vegetation index (SAVI)[J]. Remote Sensing of Environment, 1988, 25(3): 295－309.
[31] Qi J, Chehbouni A, Huete A R, et al. A modified soil adjusted vegetation index[J]. Remote Sensing of Environment, 1994, 48(2): 119－126.
[32] Mullissa A, Vollrath A, Odongo-Braun C, et al. Sentinel－1 SAR backscatter analysis ready data preparation in Google Earth Engine[J]. Remote Sensing, 2021, 13(10): 1954.
[33] Jagtap P, Shafiyoddin S. Comparative study of various single product speckle filters of SAR dataset of Sentinel－1 satellite for speckle noise reduction[J]. International Journal of Creative Research Thoughts, 2021, 9(12): c87－c95.
[34] Zheng Y, Li Z, Pan B, et al. Development of a phenology-based method for identifying sugarcane plantation areas in China using high-resolution satellite datasets[J]. Remote Sensing, 2022, 14(5): 1274.
[35] Yuan W. 30-meter resolution planting distribution dataset for sugarcane in China (2016—2020)[DS/OL]. V1. Beijing: Science Data Bank, 2023[2026－04－08]. https://doi.org/10.57760/sciencedb.IGA.00004.
[36] Yang J, Huang X. The 30 m annual land cover datasets and its dynamics in China from 1985 to 2025[DS/OL]. V1.0.5. Geneva: Zenodo, 2026[2026－04－08]. https://doi.org/10.5281/zenodo.18180184.
[37] Shen R, Pan B, Peng Q, et al. High-resolution distribution maps of single-season rice in China from 2017 to 2022[DS/OL]. V8. Beijing: Science Data Bank, 2026[2026－04－08]. https://doi.org/10.57760/sciencedb.06963.
[38] Peng Q, Shen R, Liu Y, et al. CCPD－Maize: a twenty-year dataset of maize phenology with high spatial resolution in China[DS/OL]. V3. Beijing: Science Data Bank, 2025[2026－04－08]. https://doi.org/10.57760/sciencedb.16127.
[39] Loshchilov I, Hutter F. Decoupled weight decay regularization[C]//7th International Conference on Learning Representations. New Orleans: ICLR, 2019.