引文格式：赵保睿,任玥,张敏,等. 顾及土壤温度的星载GNSS-R土壤湿度反演[J]. 地理空间信息, 2025, 23(5): 94-97. doi:10. 3969/j.issn. 1672-4623.2025.05.020

顾及土壤温度的星载GNSS-R土壤湿度反演

赵保睿[1,2]，任　玥[1*]，张　敏[2]，李　兵[2]，郭明强[2,3,4]

（1. 武汉学院，湖北 武汉 430212；2. 湖北地信科技集团股份有限公司，湖北 武汉 430074；
3. 中国地质大学（武汉）计算机学院，湖北 武汉 430074；4. 国家地理信息系统工程技术研究中心，湖北 武汉 430074）

摘　要：为提高土壤湿度的反演精度，基于全球导航定位系统反射（GNSS-R）信号数据，提出了一种随机森林支持下顾及温度的土壤湿度反演模型。结果表明，纳入温度的预测模型的平均绝对误差和均方误差分别为0.009 cm^3/cm^3和0.000 2 cm^3/cm^3，线性拟合度升至98.7%，预测的土壤湿度与真值差值的标准差约下降28.9%；纳入温度的模型可显著提高热带草原气候区域（内陆地区）的预测精度，但在湿润地区改善不明显，反演的土壤湿度空间分布与实际观测更吻合。研究结果证实了随机森林模型在土壤湿度反演领域的应用潜力以及纳入温度对提高预测模型精度的重要性。

关键词：随机森林；GNSS-R信号；土壤湿度；土壤温度；地表反射率

中图分类号：P228.4　　**文献标志码：**B　　**文章编号：**1672-4623（2025）05-0094-04

Satellite-borne GNSS-R Soil Moisture Retrieval Model Considering Soil Temperature

ZHAO Baorui[1,2], REN Yue[1], ZHANG Min[2], LI Bing[2], GUO Mingqiang[2,3,4]

(1. Wuhan University, Wuhan 430212, China;
2. Hubei Dixin Technology Group Co., Ltd., Wuhan 430074, China;
3. School of Computer Science, China University of Geosciences (Wuhan), Wuhan 430074, China;
4. National Geographic Information System Engineering Technology Research Center, Wuhan 430074, China)

Abstract: To improve the retrieval accuracy of soil moisture, based on global navigation satellite system-reflectometry (GNSS-R) signal data, we proposed a satellite-borne GNSS-R soil moisture retrieval model considering soil temperature supported by random forest. The results demonstrate that the prediction model considering soil temperature achieves a reduction in the mean absolute error and mean square error to 0.009 cm^3/cm^3 and 0.000 2 cm^3/cm^3, respectively. The linear fitting degree of it increases to 98.7%, and the standard deviation of difference between the model considering soil temperature and the true values decreases by approximately 28.9%. The prediction accuracy of model considering soil temperature is markedly enhanced in tropical savanna climate areas, but shows no significant improvement in wet regions. Additionally, the spatial distribution of soil moisture inverted by this model is more consistent with actual observations. This study confirms the potential application of random forest model in soil moisture inversion and the importance of considering soil temperature in improving prediction accuracy.

Key words: random forest, GNSS-R signal, soil moisture, soil temperature, surface reflectance

土壤湿度是陆地水循环中的关键参数，对于水文和气象过程具有深远影响[1]，是旱情预测、农业规划和气候变化模型的重要指标[2]。传统方法虽然准确，但无法捕捉大尺度土壤湿度的异质性[3]，因此迫切需要一种强大、可扩展的方法，可在更广阔区域、更高分辨率下分析土壤湿度。目前全球范围内土壤湿度的遥感监测方法主要依赖于光学遥感、微波遥感和全球导航卫星系统反射（GNSS-R）技术。光学遥感的空间分辨率较高，但易受云层和植被遮挡，难以实现全天候观测。旋风全球导航卫星系统（CYGNSS）[4]是美国宇航局专为监测热带气旋设计的卫星群[5]，能穿透云层和降水，实现全天候监测[6-8]，且每日重访周期短、数据覆盖范围广，在土壤湿度研究中具有巨大的应用潜力[9]，尤其是在微波遥感方法受限的情况下。近年来，利用CYGNSS数据进行土壤湿度反演取得了显著进展[10-18]，但仍存在不足之处：首先大多数研究未充分考虑温度对土壤湿度的影响，即使考虑了温度，往往采用线性回归，无法捕捉温度与土壤湿度之

项目来源：国家自然科学基金资助项目（41971356）；自然资源部城市国土资源监测与仿真重点实验室开放基金资助项目（KF-2022-07-001）；2019年度湖北省教育厅科学研究计划指导性资助项目（B2019334）；2018年度湖北省建设科技厅计划资助项目。
第一作者简介：赵保睿（1987—），硕士在读，高级工程师，研究方向为新型基础测绘、区域测绘基准的建设和维护等，E-mail：492318720@qq.com。
通信作者：任玥（1987—），硕士研究生，讲师，研究方向为建筑管理、建筑物变化监测、隔震结构等，E-mail：909118438@qq.com。

间的复杂非线性关系；其次许多研究依赖线性模型，未能有效处理反射率、温度等因素之间的非线性关系，限制了反演精度；再次针对不同地理区域和气候条件的适用性研究较少，区域差异性未得到充分探讨。鉴于此，本文基于CYGNSS数据，提出了一种顾及温度的土壤湿度反演模型，通过引入温度这一关键因子，重点探讨其对土壤湿度估算精度的提升作用，并通过实验对比验证了温度在提高反演方法可靠性和精度中的重要性。

1　研究数据和方法

1.1　数据来源

本文选取澳大利亚东部沿海地区作为研究区，经度范围为146.375°~153.875°，纬度范围为−34.875°~−20.125°，数据时间为2019年8—12月。CYGNSS Level 1 v3.0数据集可从PO.DAAC网站（https://podaac.jpl.nasa.gov）获取。本文采用SMAP的全球日级L3辐射计数据，来源于美国国家冰雪数据中心，空间分辨率为36 km×36 km，包含土壤湿度和质量标志等附加信息，以此作为真实值进行土壤湿度的反演和验证。本文采用欧洲中期天气预报中心第五代再分析数据（ERA5）数据集提供的Soil Temperature Level1数据。该数据是表征土壤浅表层（0~7 cm）热力状态的关键参数，土壤温度在各层中心位置计算，以K为单位；空间分辨率为0.25°×0.25°、时间分辨率为1 h。Soil Temperature Level1基于热扩散方程，通过能量守恒的隐式算法进行计算，确保在表面强迫周期（0.5~730 d）内幅度误差小于20%，相位误差小于5%，具有较高的精度[19]。模型假设土壤体积热容量为$2.19\times10^{6}\ \mathrm{J\cdot m^{-3}\cdot K^{-1}}$，热导率则根据土壤含水量采用Peters-Lidard方法计算；同时假设底部边界条件无热通量，以保证深层土壤温度波幅衰减至可忽略程度。

1.2　基于CYGNSS数据反演土壤湿度

在相干假设中，入射至地表的GNSS信号在陆地表面仅发生相干反射。去除双基地雷达几何和观测系统硬件的影响，地表反射率可定义为：

$$S=\frac{(4\pi)^2P_{\mathrm{coh}}(R_{\mathrm{ts}}+R_{\mathrm{rs}})^2}{\lambda^2G_{\mathrm{r}}P_{\mathrm{t}}G_{\mathrm{t}}}$$

式中，P_{coh}为延迟多普勒图的峰值功率；P_{t}为GNSS卫星右旋圆极化导航信号的发射功率；G_{t}、G_{r}为发射机、接收机天线增益；λ为发射信号的载波波长；R_{ts}、R_{rs}为发射机、接收机到地表反射点的距离。

基于地表反射率和温度，利用随机森林方法预测土壤湿度，以SMAP土壤湿度为真实值对模型进行验证。具体步骤包括：①数据预处理与过滤，先通过空间匹配排除海洋区域数据，以提高数据质量和一致性，再对所有数据进行时间和空间校准，以确保反演数据处于同一时空尺度；②特征选择与数据分割，以经度、纬度、地表反射率和温度为输入特征，土壤湿度为目标变量，将数据集分割为训练集（前70%，2019年8—11月中旬）和测试集（后30%，2019年11月下旬—12月底），确保了模型的泛化能力；③随机森林模型训练，利用随机森林回归器进行土壤湿度反演的模型训练，模型参数设置为100棵决策树以平衡计算开销与预测性能，最大深度设置为20以防止过拟合，最小分裂样本数设置为5，最小叶子节点样本数设置为2，并采用平方根方式选择分裂特征，再利用自助采样法构建训练数据集，确保模型的随机性和稳定性，最后通过5折交叉验证优化超参数，确保模型在训练集和测试集上均表现出良好的泛化能力，这些参数经过多次实验验证，可保证模型在处理输入特征复杂性时的鲁棒性和准确性；④模型验证与评估，将训练好的随机森林模型应用于测试集数据，并采用平均绝对误差（MAE）、均方误差（MSE）和决定系数（R^2）评估模型的预测性能；⑤模型优化与对比分析，为进一步评估模型的表现，将不包含温度变量的模型与包含温度的模型进行对比分析，以量化温度在土壤湿度反演中的贡献。

2　研究结果与分析

2.1　基于地表反射率反演土壤湿度

本文首先采用传统输入数据（地表反射率）进行训练，通过大量试验得出了最优参数策略：决策树数量设置为100，以确保模型具备足够的能力捕捉数据中的复杂关系；随机数种子设置为42；其他参数均采用随机森林算法的默认设置。基于上述参数策略，以70%的数据为训练集，30%的数据为验证集，训练得到预测模型；再以SMAP土壤湿度为真实值，对模型进行验证。验证集的预测结果见图1，可以看出，实验区土壤湿度大多维持0.2 $\mathrm{cm^3/cm^3}$以下，预测土壤湿度和真实值基本分布于$y=x$直线附近；土壤湿度较低时预测精度较高，土壤湿度超过0.2 $\mathrm{cm^3/cm^3}$时，离散程度开始增大，而超过0.4 $\mathrm{cm^3/cm^3}$后，离散程度又呈减小趋势；模型的MAE、MSE分别为0.019 $\mathrm{cm^3/cm^3}$、0.000 7 $\mathrm{cm^3/cm^3}$；线性拟合的公式为$y=0.97x+0.21$，拟合度为96.1%。

验证集的土壤湿度反演误差分布情况见图2，可以看出，误差基本呈正态分布，无明显的系统误差，说

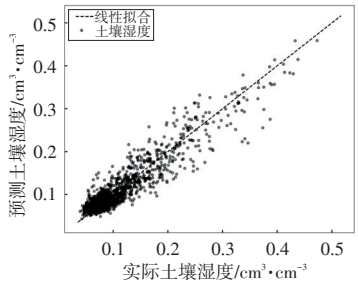


图1 基于地表反射率预测土壤湿度与真实值对比

明训练的模型比较准确；误差在±0.2 cm³/cm³以内的占比为71.25%，在±0.2~±0.5 cm³/cm³之间的占比为23.16%，大于0.5 cm³/cm³的占比为5.59%。综合来看，虽然模型预测精度较高，但在特定区域仍有优化空间。

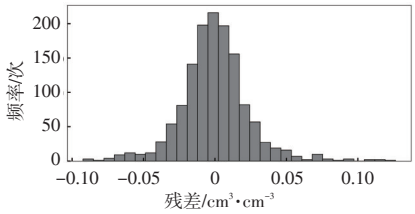


图2 基于地表反射率反演土壤湿度误差分布

利用训练成熟的模型，基于验证集数据对土壤湿度进行预测，并以格网点为基础对预测结果进行平均，与真实值进行对比，结果见图3，可以看出，预测值和真实值的一致性较高，且随土壤湿度的增加而降低；实验区土壤湿度由西向东、由北向南、由内陆向沿海逐渐增加，符合当地的自然规律。

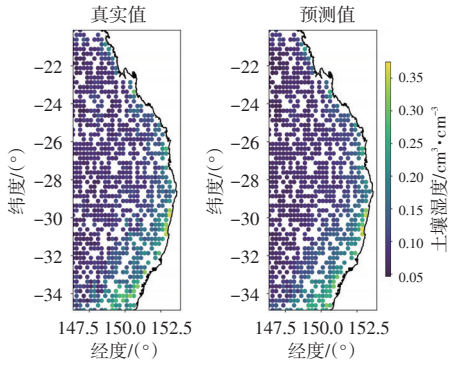


图3 预测的土壤湿度与真实值的空间对比

由于澳洲东部沿海区域属热带雨林和亚热带湿润气候，气候湿润多雨；而内陆区域是热带草原气候。本文以离海岸距离表征干旱程度（离海岸越远、气候越干燥），研究预测误差与气候湿润程度的关系。验证集预测误差结果见图4，可以看出，在热带雨林和亚热带湿润气候区域（离海岸2°以内），模型的预测效果较好，原因在于该区域土壤湿度主要受降水控制，温度的影响较小；而在内陆热带草原气候区域，气候干燥且温度波动较大，模型未能充分反映温度对土壤湿度的影响，导致预测精度较低，尤其是在干旱和高温季节，模型偏差较明显。

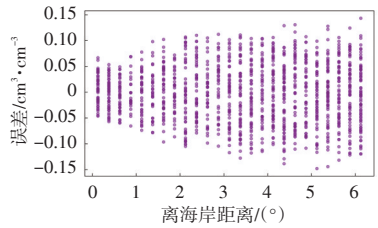


图4 预测误差与离海岸距离的关系

2.2 顾及温度的土壤湿度反演

土壤湿度与温度联系密切，为提高土壤湿度的反演精度，本文在模型训练时考虑了温度参数，构建了顾及温度的预测模型，验证集的预测结果见图5，可以看出，顾及温度的预测模型的预测精度明显提高，预测值与真实值的分布更加接近 $y=x$ 直线；在土壤湿度超过0.3 cm³/cm³时，预测值与真实值的一致性出现下降，比未考虑温度的模型提高了0.1 cm³/cm³普适性得到了明显提高；模型的MAE、MSE分别为0.009 cm³/cm³、0.000 2 cm³/cm³比未考虑温度的模型分别降低了52.6%和71.4%；线性拟合公式为 $y=1.01x+0.08$，拟合度为98.7%，比未考虑温度的模型提高了2.6%

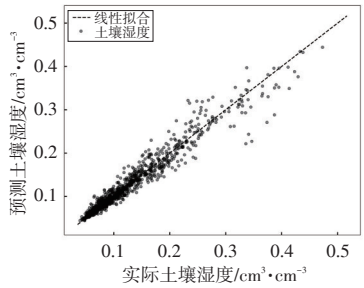


图5 基于地表反射率和温度预测土壤湿度与真实值对比

顾及温度的模型在验证集的预测残差分布（图6）更符合正态分布特征，预测误差明显降低；误差在±0.2 cm³/cm³以内的占比为88.68%，在±0.2~±0.5 cm³/cm³之间的占比为9.36%，大于0.5 cm³/cm³的占比为1.96%，比未考虑温度的模型降低了35.06%，模型精度提升明显。

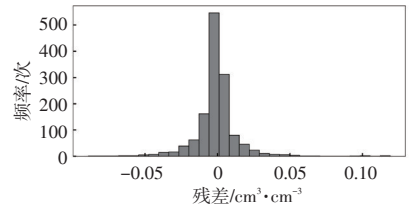


图6 基于地表反射率和温度反演土壤湿度误差分布

利用顾及温度的预测模型，基于验证集数据对土壤湿度进行预测，并以格网点为基础对预测结果进行平均，与真实值进行对比，结果见图7，可以看出，土壤湿度的空间分布特点未发生明显变化，但在内陆区域预测值和真实值的符合度有一定提高。

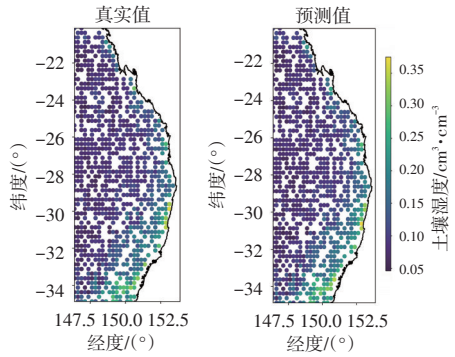


图7　基于地表反射率和温度预测的土壤湿度与真实值的空间对比

为分析引入温度后，模型对不同气候区域的改善程度，本文对顾及温度的预测误差与气候湿润程度的关系进行分析，结果见图8，可以看出，引入温度后模型在湿润地区的改善不明显，但在热带草原气候区域（内陆地区）的预测精度显著提高。由于内陆区域气候干燥，土壤湿度对温度更敏感，温度上升导致土壤水分迅速蒸发，预测模型通过温度捕捉到了这些变化，因此能更准确地反映干旱区域的土壤湿度动态变化。在沿海湿润区域，温度与土壤湿度之间可能出现解耦现象，导致温度对湿度预测的贡献有限，降雨仍是主要驱动力，模型预测精度没有明显变化。值得注意的是，温度与土壤湿度之间的关系在不同气候区域存在差异，温度信息在不同区域和气候条件下对湿度反演结果的提升存在一定的限制性，这也与土壤湿度遥感领域的已有结论一致。未来研究可进一步结合降水量、植被指数等其他环境变量来优化模型在不同区域的适用性。

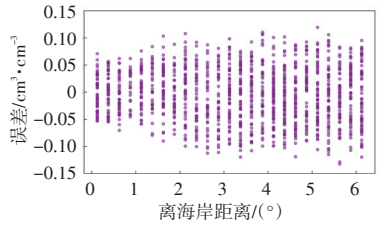


图8　引入温度后预测误差与离岸距离的关系

为量化纳入温度前后模型的性能差异，本文分别对两种模型的预测结果与真实值的差值进行统计，结果见表1，顾及温度的模型在整体性能上显著提升，差值均值由0.013 cm^3/cm^3降至0.007 cm^3/cm^3，约减少了46.2%，表明预测误差得到明显控制，标准差由0.045 cm^3/cm^3降至0.032 cm^3/cm^3，约下降了28.9%，表明预测结果的稳定性和一致性得到了提升；虽然最大、最小值差异略有缩小，但进一步证明了顾及温度的模型能更有效地捕捉温度对土壤湿度的非线性影响，减少系统误差和离散程度。整体来看，温度的引入使模型预测精度和稳定性得到优化，验证了其在土壤湿度反演中的重要性。

表1　模型预测土壤湿度与真实值的差值统计结果/$cm^3\cdot cm^{-3}$

指标	未顾及温度的模型	顾及温度的模型
最大值	0.121	0.101
最小值	0.119	0.109
平均值	0.013	0.007
标准差	0.045	0.032

3　结　论

本文以澳大利亚东部地区为研究对象，提出了一种基于随机森林算法顾及温度的反演模型；并对比了纳入温度变量前后的反演结果，得出如下结论：

1）在未顾及温度的模型中，当土壤湿度低于0.2 cm^3/cm^3时，预测精度较高，模型的MAE、MSE分别为0.019 cm^3/cm^3和0.000 7 cm^3/cm^3，线性拟合度为96.1%。

2）在顾及温度的模型中，预测值与真实值的分布更加接近 $y=x$ 直线。当土壤湿度超过0.3 cm^3/cm^3时，预测精度出现下降，比未顾及温度的模型提高了0.1 cm^3/cm^3，普适性得到了明显提高。模型的MAE、MSE分别为0.009 cm^3/cm^3和0.000 2 cm^3/cm^3，比未顾及温度的模型分别降低了52.6%和71.4%。顾及温度的模型预测的土壤湿度与真实值差值的标准差约下降28.9%，模型在湿润地区改善不明显，但在热带草原气候区域（内陆地区）的预测精度显著提高。

3）土壤湿度的空间分布特点表明，湿度由西向东、由北向南、由内陆向沿海逐渐增加，与当地自然地理规律相符。顾及温度的模型在空间分布预测中更准确，预测值与真实值的符合度更高，误差分布更趋于正态，更加贴近实际的空间变异性。

参考文献

[1] 陈立文,陈晨,唐晓玲,等. 基于RS的淮河流域土壤含水量估算方法[J]. 地理空间信息,2016,14(11):97-102

[2] 焦振航,舒红,吴凯,等. 水文模型的LAI优化与单网格土壤湿度同化[J]. 地理空间信息,2019,17(9):68-72

[3] 张志刚,梁月吉,任超,等. 小波基在GPS-IR反演土壤湿度中优选方法的研究[J]. 地球物理学进展,2020,35(1):72-78

[4] Yueh S H, Shah R, Chaubell M J, et al. A Semiempirical Modeling of Soil Moisture, Vegetation, and Surface Roughness Impact on CYGNSS Reflectometry Data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 60: 1-17

[5] 王俊栋,孙志刚,杨婷,等. CYGNSS土壤盐分反演方法研究:以黄河三角洲为例[J]. 遥感学报,2023,27(2):351-362

[6] 汪俊涛,朱勇超,高飞,等. 基于CYGNSS数据的星载GNSS-R水体分布探测方法研究[J]. 大地测量与地球动力学,2021,41(8):777-782

（下转第101页）