大河湾试验区土壤有机质遥感反演实验报告（02-08）

1. **实验背景**

土壤有机质含量是衡量土壤肥力的重要指标。通常情况下，光谱各波段的反射率随有机质含量的升高而不断降低，二者表现出较强的负相关性，因此可以通过遥感影像对土壤有机质含量进行大空间尺度反演，实现宏观监测，为农业发展提供数据支撑。

1. **实验数据**

本次实验可选用的遥感数据为GF-1（2021年6、8、9月）、Landsat8-OLI（2019年10月）和Sentinel2（2021年4月）共3类多光谱影像，影像的选取原则为“近期、裸土、少云、无雪及霜冻”。东北地区春耕时间较晚，多为四五月份，自6月起地表开始被植被覆盖，因此GF-1多光谱3景影像均不适用于土壤有机质反演。本次实验最终使用了Landsat8卫星和Sentinel2卫星的多光谱数据，二者均经过了大气校正，Sentinel2影像为经过了正射校正和大气校正的L2A级数据。

本次实验采用的实际采样数据为大河湾农场地区共185个采样点的有机质监测结果。

**3. 实验建模方法**

3.1 相关性分析

对遥感影像和实测数据进行相关性分析的目的是找出研究对象（有机质含量）的敏感波段，此步骤对于高精度反演来说至关重要。相关性通过相关系数体现，计算方法为：



式中和分别为影像反射率及反射率的平均值，和分别为土壤有机质实测值及其平均值。为正则表示二者成正相关，为负则表示负相关，绝对值越大则相关性越强。

3.2 模型选择

有机质含量的增加会使土壤在各波段的反射率均降低，因此本实验采用一元及多元线性回归模型，训练集和测试集按总采样量进行4：1分配，设置变量方差显著性水平应按5%作为剔除的阈值。

3.3 模型检验

模型的检验包含两个方面，分别是稳定性检验和准确度检验（也称预测精度）。稳定性的衡量指标时判定系数，判定系数介于0~1之间，数值越大，则表明模型越稳定，出错的可能性越小。准确度的衡量指标是均方根误差RMSE，这个值越小表示模型预测的越准确。

**4. 实验结果分析**

4.1 相关性分析结果

根据相关系数计算结果，Landsat8多光谱影像和Sentinel2多光谱影像土壤表面反射率与SOM实测值之间的相关性均不高，显著性检验结果也表明样本与总体之间的关系不具备足够的代表性。各波段相关系数见表1：

**表1 Landsat8及Sentinel2反射率与实测值相关系数**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 波段号  数据源 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Landsat8 | 0.22 | 0.24 | 0.21 | 0.24 | -0.25 | 0.19 | 0.17 | / | / | / | / | / |
| Sentinel2 | 0.06 | 0.03 | -0.05 | -0.11 | -0.1 | -0.1 | -0.1 | -0.06 | -0.1 | 0 | 0.05 | 0.01 |

4.2 模型精度

根据相关性计算可知，两组数据中Landsat8多光谱反射率与实测值取负值相关性更高，因此模型使用Landsat8影像第五波段参与反演，经一元线性回归后得到反演公式如下：



从该结果表达式看，整体上能够体现出土壤有机质含量越高反射率越低的性质，但损失函数并没有收敛到可接受范围，经计算，与采样数据相对应的的均方根误差RMSE为2.44，判定系数也仅为0.153，说明模型的系数需要调整。

多元线性回归的原理与一元类似，结果也会遇到相似的问题。

**5. 阶段小结**

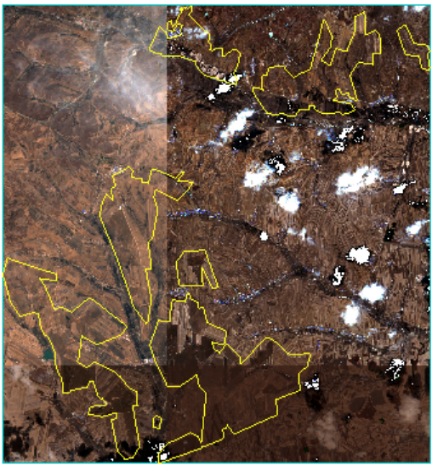


图1 土壤有机质反演结果与sentinel2高分辨率真彩色影像对比

从反演结果中可以看出，有机质的空间分布情况与目视判读情况吻合，说明该模型在校正参数后可以用于反演土壤有机质含量。由于采样结果与多个传感器观测结果之间均没有良好的相关性，所以下一步需解决该问题。