引文格式：艾小童. 基于ELM模型的极化SAR土壤水分降尺度研究[J]. 地理空间信息, 2023, 21(7): 11-15. doi:10. 3969/j.issn. 1672-4623.2023.07.003

基于ELM模型的极化SAR土壤水分降尺度研究

艾小童[1]
（1. 武汉大学 资源与环境科学学院，湖北 武汉 430079）

摘　要：为实现土壤水分实时和广域监测，解决被动微波遥感得到的土壤水分数据分辨率低的问题，利用极限学习机（ELM）模型建立极化SAR数据与土壤水分的反演模型，再选择最合适的映射函数实现SMAP土壤水分数据的降尺度研究。结果表明，以三角函数为ELM模型的映射函数进行土壤水分降尺度研究的误差最小，并可取得高分辨率土壤水分数据。将该研究应用于正蓝旗区域发现，各月份土壤水分空间分布变化较大，4月、6月北部草原降水量较高，而8月、10月南部草原降水量较高。

关键词：土壤水分；极化SAR；ELM；遥感反演
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2023）07-0011-05

Soil Moisture Downscaling Based on Extreme Learning Machine Model Using Polarimetric SAR Images

AI Xiaotong[1]
(1. School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China)

Abstract: In order to realize the real-time and expansive monitoring of soil moisture and solve the problem of low resolution of soil moisture data obtained by passive microwave remote sensing, we used extreme learning machine(ELM) model to establish the inversion model between polarimetric SAR data and soil moisture, and selected the most appropriate mapping function to realize the downscaling of SMAP soil moisture data. The results showed that using trigonometric function as the mapping function of ELM model to study soil moisture downscaling had the least error, and could obtain high-resolution soil moisture data. Applying this study to Zhenglan Banner area, we found that the spatial distribution of soil moisture changed greatly in different months, and the precipitation of northern grassland was higher in April and June, while that of southern grassland was higher in August and October.

Key words: soil moisture, polarimetric SAR, ELM, remote sensing inversion

土壤水分不仅影响植物生长，还影响陆气之间物质与能量的交换[1]，因此土壤水分监测在农业、气象和环境监测等领域具有广泛应用。相较于传统野外调查，基于遥感影像的土壤水分监测方式具有监测范围广、更新周期短和信息量大的优点[2]，其中微波遥感不受光线、气候和云雾限制，且与地表土壤物理属性具有强相关性，被广泛应用于土壤含水量反演研究中[3]。然而，被动微波遥感监测到的土壤水分数据空间分辨率较低，限制了更广泛的水文计算和小尺度的土壤水分监测[4]。利用主动微波遥感手段——合成孔径雷达（SAR）数据与被动微波遥感监测到的土壤水分数据建立土壤水分反演模型，进行土壤水分数据的降尺度研究，即可获得高分辨率的土壤水分数据[5]。目前，极化SAR土壤水分反演的研究非常丰富，但几乎没有论文利用极化SAR数据进行土壤水分的降尺度研究[6-7]。

本文以内蒙古自治区锡林郭勒盟正蓝旗区域为研究对象，通过Sentinel-1和Sentinel-2卫星影像获取后向散射系数、微波入射角、NDWI等数据；再通过水云模型计算裸土后向散射系数，通过Dubois模型计算土壤介电常数，通过AIEM模型计算地表粗糙度；然后以上述指标为输入样本，以具有实时性和全球性的SMAP土壤水分数据为输出样本，通过极限学习机（ELM）模型建立土壤水分反演模型，从而获取研究区域2021年4月、6月、8月、10月的高分辨率土壤水分数据；最后对研究区域内土壤水分数据分布和时空变化情况进行对比分析。

1　研究区概况与数据来源

1.1　研究区概况

内蒙古自治区锡林郭勒盟正蓝旗，位于内蒙古自

收稿日期：2022-05-12；修回日期：2022-07-14。
项目来源：国家重点研发计划资助项目（2018YFD1100801-01）。
作者简介：艾小童（1999—），硕士研究生，主要从事遥感反演、摄影测量、机器学习等方面的研究，E-mail：2017301610022@whu.edu.cn。

治区中部，锡林郭勒草原东南边缘；属中温带大陆性气候，年均气温为1.5℃；年降雨量为365.1 mm，主要集中在夏季，冬天有180 d的冰雪期。

1.2 SMAP数据

本文选取NASA-USDA增强型SMAP全球土壤水分数据作为反演模型的输出数据，空间分辨率为10 km，时间分辨率为1 d。SMAP土壤水分数据是由SMAP卫星搭载的被动辐射计测量的地球表面亮度和温度数据反演得到的土壤水分观测结果。NASA-USDA增强型SMAP全球土壤水分数据是利用一维卡尔曼滤波器将SMAP土壤水分数据集成到改进的双层帕尔默模型中生成的，可在谷歌公司的开放式云计算GEE平台中下载，具体内容见表1。

表1 NASA-USDA增强型SMAP数据包含的内容

名称	含义	最小值	最大值
ssm	表层土壤水分/mm	0	25.39
susm	次表层土壤水分/mm	0	274.6
smp	土壤水分/%	0	1
ssma	表层土壤水分异常	-4	4
susma	次表层土壤水分异常	-4	4

1.3 Sentinel数据

Sentinel-1搭载了C波段SAR，包含地面距离探测（GRD）场景，有10 m、25 m和40 m三种分辨率，VV、HH、VV+VH、HH+HV四种波段组合[8-9]。本文主要采用VV波段，即垂直发射垂直接收雷达信号的波段，植被层对电磁波散射和吸收的影响较小，更适用于土壤水分反演。利用GEE平台对Sentinel-1影像进行预处理，主要包括利用轨道文件更新轨道元数据、消除GRD边界噪声、消除热噪声、辐射测量校准、正射校正。本文以Sentinel-2卫星影像为光学影像数据，包括10 m、20 m和60 m三种分辨率。利用GEE平台获取Sentinel-2 L2A影像数据，并对其进行去云处理，基于影像中的近红外波段和红光波段计算NDWI用于植被水分的反演[10]。

2 研究方法

由于大部分地表有植被覆盖，且往往凹凸不平，这些都会对微波遥感造成影响；而土壤介电性质与土壤体积含水量之间联系密切，且通过土壤介电性质来反演土壤水分的方法已得到广泛应用，因此本文利用土壤介电常数来提高反演模型的相关性。本文通过水云模型、Dubois模型和AIEM模型分别计算裸土的后向散射系数、土壤介电常数和地表粗糙度，并将其作为ELM模型的输入样本，SMAP数据作为输出样本，反演得到土壤水分数据。整体流程见图1。

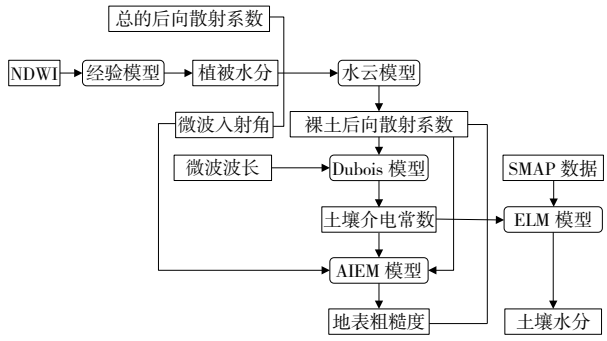


图1 整体流程图

2.1 水云模型

极化SAR土壤水分反演往往通过雷达后向散射系数与土壤水分之间的关系建立反演模型；但由于植被会影响电磁波的散射和吸收，通常土壤水分反演模型易受植被覆盖的影响。水云模型能有效避免植被影响，将植被覆盖区域的后向散射定义为植被自身反射微波得到的体散射与经过植被二次反射后的地表散射的和[11]，表达式为：

$$\begin{cases}\sigma^0_{can}(\theta)=\sigma^0_{veg}(\theta)+\gamma^2(\theta)\times\sigma^0_{soil}(\theta)\\ \sigma^0_{veg}(\theta)=A\times m_{veg}\times\cos\theta\times(1-\gamma^2(\theta))\\ \gamma^2(\theta)=\exp(-2\times B\times m_{veg}\times\sec\theta)\end{cases}\quad(1)$$

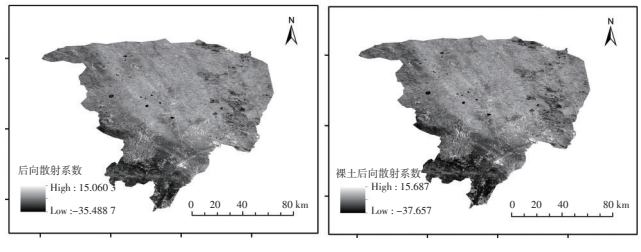
式中，$\sigma^0_{can}(\theta)$、$\sigma^0_{soil}(\theta)$、$\sigma^0_{veg}(\theta)$分别为草地、裸土和植被的后向散射系数；A、B为经验系数；m_{veg}为植被水分；θ为微波入射角；$\gamma^2(\theta)$为二次衰减因子。

草地的后向散射系数和微波入射角可从处理好的Sentinel-1影像中直接获取，植被水分可通过NDWI反演得到，经验系数则依赖于植被类型。通过经验模型利用NDWI对植被水分进行反演，表达式为[12]：

$$m_{veg}=1.44\times NDWI^2+1.36\times NDWI+0.34\quad(2)$$

不同植被覆盖下的经验系数不同，牧地的水云模型经验系数A=0.001 4、B=0.084[13]。

利用ArcGIS10.6中的Raster Calculator工具对整个研究区域进行水云模型计算，结果见图2。



a 总后向散射系数 b 裸土后向散射系数

图2 后向散射系数对比图

2.2 Dubois模型

土壤水分能影响土壤介电常数，因此可利用土壤介电常数反演土壤水分。Dubois模型是一种半经验模

型，在土壤介电常数与裸土后向散射系数之间建立联系，表达式为[14]：

$$\begin{cases}\sigma_{\mathrm{hh}}^{0}=10^{-2.75}\dfrac{\cos^{1.5}\theta}{\sin^{5}\theta}10^{0.028\varepsilon_r\times\tan\theta}(k\times s\times\sin\theta)^{1.4}\lambda^{0.7}\\ \sigma_{\mathrm{vv}}^{0}=10^{-2.35}\dfrac{\cos^{3}\theta}{\sin^{3}\theta}10^{0.046\varepsilon_r\times\tan\theta}(k\times s\times\sin\theta)^{1.1}\lambda^{0.7}\end{cases}\tag{3}$$

式中，σ_{hh}^{0}、σ_{vv}^{0}分别为HH和VV波段的裸土后向散射系数；k为入射波波数；λ为微波波长；s为地表参数；ε_r为土壤介电常数。

2.3　AIEM模型

地表粗糙度能影响雷达后向散射系数，因此土壤水分反演常用到地表粗糙度数据。积分方程模型（IEM）是基于电磁波辐射传输理论建立的地表散射模型，可反映地表粗糙度的真实情况[15]。AIEM模型则是IEM的改进，对地表粗糙度的模拟更准确[16]，表达式为：

$$\begin{cases}I_{\mathrm{pq}}^{n}=(2k\cos\theta)^{n}\times f_{\mathrm{pq}}\times e^{-S^2(k\cos\theta)^2}+\\ \dfrac{(k\cos\theta)^{n}\left[F_{\mathrm{pq}}(-k\sin\theta,0)+F_{\mathrm{pq}}(k\sin\theta,0)\right]}{2}\\ \sigma_{\mathrm{pq}}^{0}=\dfrac{k^2}{2}e^{-2S^2(k\cos\theta)^2}\sum_{n=1}^{\infty}S^{2n}\left|I_{\mathrm{pq}}^{n}\right|^{2}\dfrac{W^{n}(-2k\sin\theta,0)}{n!}\end{cases}\tag{4}$$

式中，σ_{pq}^{0}为裸土后向散射系数；pq为极化方式；S为均方根高度；I_{pq}^{n}为中间变量；$W^{n}(-2k\sin\theta,0)$为地表粗糙度谱；f_{pq}为基尔霍夫场系数；F_{pq}为补偿场系数。

F_{pq}、f_{pq}的求解公式为：

$$\begin{cases}F_{\mathrm{vv}}(-k\sin\theta,0)+F_{\mathrm{vv}}(k\sin\theta,0)=\dfrac{2\sin^2\theta(1+R_{\mathrm{vv}})^2}{\cos\theta}\times\\ \left[1-\dfrac{1}{\varepsilon_r}+\dfrac{\varepsilon_r-\sin^2\theta-\varepsilon_r\cos^2\theta}{\varepsilon_r^2\cos^2\theta}\right]\\ F_{\mathrm{hh}}(-k\sin\theta,0)+F_{\mathrm{hh}}(k\sin\theta,0)=\\ -\dfrac{2\sin^2\theta(1+R_{\mathrm{hh}})^2}{\cos^3\theta}(\varepsilon_r-1)\end{cases}\tag{5}$$

$$\begin{cases}f_{\mathrm{vv}}=\dfrac{2(\varepsilon_r\cos\theta-\sqrt{\varepsilon_r-\sin^2\theta})}{\cos\theta(\varepsilon_r\cos\theta+\sqrt{\varepsilon_r-\sin^2\theta})}\\ f_{\mathrm{hh}}=-\dfrac{2(\cos\theta-\sqrt{\varepsilon_r-\sin^2\theta})}{\cos\theta(\cos\theta+\sqrt{\varepsilon_r-\sin^2\theta})}\end{cases}\tag{6}$$

Baghdadi N[17]等通过大量数据，基于IEM模型实践模拟得到不同极化方式的长度定标公式，构建均方根高度反演地表粗糙度的半经验模型，即

$$\begin{cases}l_{\mathrm{opt}}(\mathrm{HH},S,\theta)=0.162+3.006(\sin 1.23\theta)^{-1.494S}\\ l_{\mathrm{opt}}(\mathrm{VV},S,\theta)=1.281+0.134(\sin 0.19\theta)^{-1.59S}\end{cases}\tag{7}$$

式中，l_{opt}为地表粗糙度参数定标后的长度。

2.4　ELM模型

ELM模型是一种基于前馈神经网络构建的机器学习系统或方法，包括输入层、隐含层和输出层，通过线性方程组求解各层之间的权值[18]。由于ELM模型无需迭代运算，因此学习训练速度快，适用于大范围土壤水分反演监测工作，且避免了过度拟合等问题。本文以裸土后向散射系数、土壤介电常数和地表粗糙度为模型的输入样本，算法结构见图3。

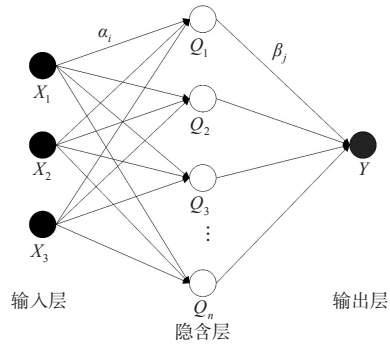


图3　ELM模型算法结构图

隐含层的输出模型为：

$$f_L=\sum_{i=1}^{l}\beta_i h_i(x)=h(x)\beta\tag{8}$$

$$h(x)=G(a_i,b_i,x)\tag{9}$$

式中，x为ELM模型的输入值；β为输出权重；$h(x)$为映射函数；a_i为输入层与隐含层之间的权重；b_i为节点参数。

ELM模型构建的具体过程为：

1）初始化ELM参数。

$$\boldsymbol{A}=\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\a_{31}&a_{32}&\cdots&a_{3n}\end{bmatrix}\quad\boldsymbol{K}=\begin{bmatrix}K_1\\K_2\\K_3\end{bmatrix}\tag{10}$$

式中，$\boldsymbol{A}$为输入层与隐含层权值矩阵；$\boldsymbol{K}$为隐含层阈值矩阵。

2）计算输出权重。

$$\boldsymbol{B}=\begin{bmatrix}G(a_1,b_1,x_1)&G(a_2,b_2,x_1)&\cdots&G(a_n,b_n,x_1)\\G(a_1,b_1,x_2)&G(a_2,b_2,x_2)&\cdots&G(a_n,b_n,x_2)\\G(a_1,b_1,x_3)&G(a_2,b_2,x_3)&\cdots&G(a_n,b_n,x_3)\end{bmatrix}\tag{11}$$

$$\beta=\boldsymbol{B}^{+}\times Y$$

式中，$\boldsymbol{B}$为隐含层的输出矩阵；$\boldsymbol{B}^{+}$为$\boldsymbol{B}$的广义逆矩阵；Y为输出结果；$G()$为映射函数。

本文采用的映射函数及其解析式见表2。

3　研究结果与分析

3.1　各映射函数的精度评价

研究区内SMAP土壤水分数据共占有111个栅格，即每期能得到111组样本数据。为减少相同栅格下不同区域真实土壤水分的差异，将相同栅格下的裸

表2 本文采用的映射函数及其解析式

函数名称	函数解析式
Sigmoid函数	$G(a_i, b_i, x)=\frac{1}{1+\exp(a\times x+b)}$
径向基函数	$G(a_i, b_i, x)=\exp(-b\left\|x-a\right\|_2)$
高斯函数	$G(a_i, b_i, x)=\sqrt{\left\|x-a\right\|_2+b^2}$
三角函数	$G(a_i, b_i, x)=\cos(a\times x+b)$
双曲正弦函数	$G(a_i, b_i, x)=\frac{1-\exp(a\times x+b)}{1+\exp(a\times x+b)}$
对数函数	$G(a_i, b_i, x)=1+\ln(a\times x+b)$

土后向散射系数、土壤介电常数和地表粗糙度取平均值，再与该栅格对应的土壤水分进行反演分析。对样本数据随机排序，以81组数据为训练数据，30组数据为测试数据进行实验。隐含层神经元个数和映射函数均能影响ELM模型反演结果的精度，隐含层神经元个数过少会导致拟合精度不够，而个数过多则会造成过度拟合，影响模型反演的效率。本文利用Matlab 2020a软件编写ELM函数，再分别采用6种映射函数在不同隐含层神经元数目下进行模型精度分析实验，以误差平方和为精度度量指标，找出各映射函数适应最佳的神经元数目，对比不同映射函数的精度。实验结果见图4，可以看出，不同映射函数适应的最佳隐含层神经元数目不同，且随着隐含层神经元数目的增加，反演精度变化情况也不同，Sigmoid函数、径向基函数和三角函数的反演精度均在隐含层神经元数目增至一定值后趋于稳定；而高斯函数、双曲正弦函数和对数函数则在神经元数目增至一定值后开始上下波动。由径向基函数对应的曲线可知，曲线稳定后，其反演精度反而低于稳定前，说明并非隐含层神经元数目越多，反演精度越高；在某些情况下，其反演精度反而越低。对比6种映射函数各自最佳隐含层神经元数目下的反演精度可知，三角函数的反演精度明显优于其他函数。综上所述，本文选取ELM模型的隐含层神经元数目为40，映射函数为三角函数。

3.2 土壤水分降尺度结果分析

由于正蓝旗在11月—次年3月会出现降雪，通过Sentinel-1、Sentinel-2遥感影像获取该时期土壤水分反演模型输入样本数据较困难，因此本文选择正蓝旗4月、6月、8月、10月的遥感影像和SMAP数据进行土壤水分降尺度分析研究，结果见图5，可以看出，4月正蓝旗土壤水分反演精度明显低于其他月份，这是因为4月正蓝旗草地部分区域仍有积雪存在，对输入样本数据精度造成了很大影响。

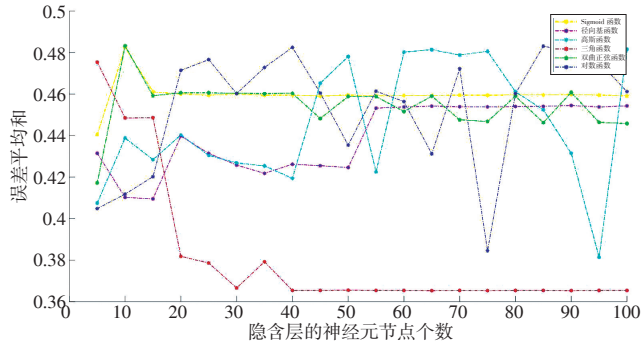


图4 不同数目隐含层神经元下各映射函数反演精度对比

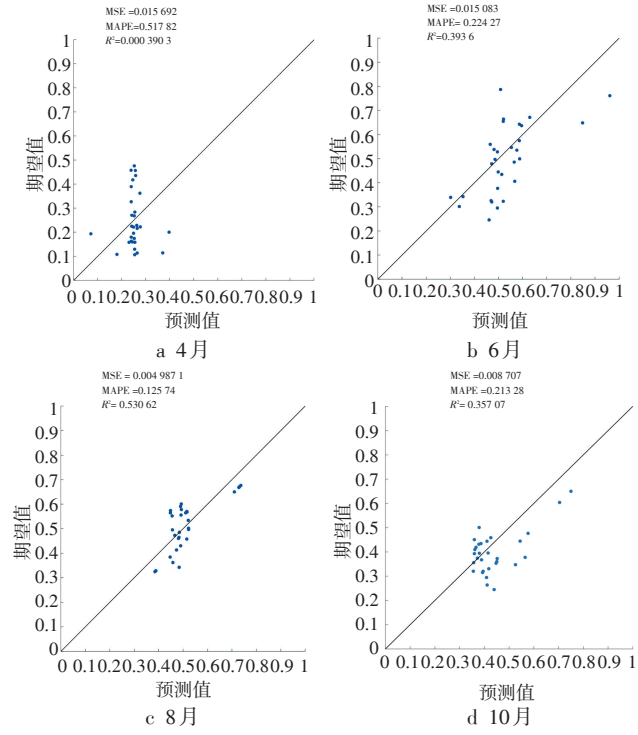


图5 各月份正蓝旗土壤水分的反演精度对比

降尺度前后正蓝旗区域土壤水分分布对比见图6，降尺度后能较好地将土壤水分数据的分辨率由10 km提高到10 m，可在更广泛的水文计算和小尺度土壤水分研究中得到应用。

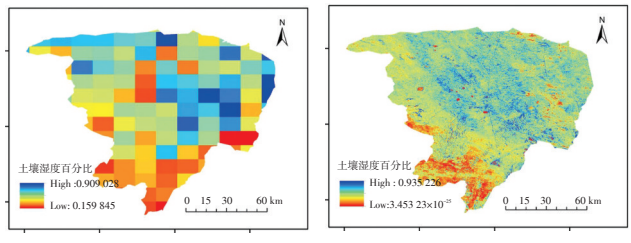
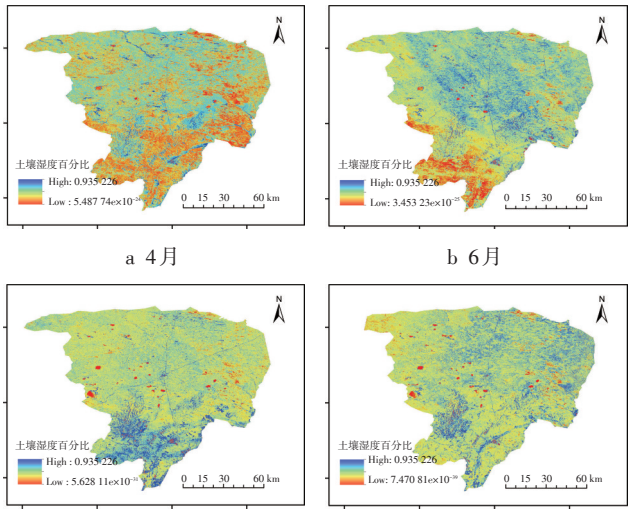


图6 降尺度前后正蓝旗土壤水分分布对比图

4月、6月、8月、10月正蓝旗区域土壤水分分布见图7，由于不同时间不同地区的降水量有所差异，因此正蓝旗区域土壤水分的空间分布随时间的变化幅度较大，正蓝旗夏季降水比春、秋季多，6月、8月的土壤水分比4月、10月高。由于4月、6月正蓝旗北部土壤湿润而南部土壤较干旱，8月、10月南部土壤

则比北部土壤更湿润，因此正蓝旗区域北部在4月和6月降水量更多，南部在8月、10月降水量更多。



a 4月　　b 6月

c 8月　　d 10月

图7　各月份正蓝旗区域土壤水分分布图

研究不同时期正蓝旗区域土壤水分分布变化情况，有助于当地农业灌溉方式以及培养农作物和牧草种类的选择，对当地农牧业发展具有一定的积极作用。

4　结　语

本文利用极化SAR对SMAP土壤水分数据进行降尺度研究，归纳整理了基于ELM模型的极化SAR土壤水分降尺度分析的技术要点；并将该方法应用于正蓝旗区域，对该区域2021年土壤水分的反演精度和时空变化进行了分析，得出如下结论：

1）在土壤水分降尺度研究中，ELM模型的映射函数为Sigmoid函数、径向基函数和三角函数时隐含层神经元数目增至一定值后会趋于稳定，且映射函数为三角函数时模型的反演精度明显优于其他函数。

2）ELM模型虽能高效反演大范围的土壤水分数据，但反演精度有限，如何高效率高精度反演土壤水分，需要进一步研究。

3）不同时期正蓝旗区域土壤水分空间分布差异较大，4月、6月北部降水量更多，8月、10月南部降水量更多。

4）土壤水分除与裸土后向散射系数、土壤介电常数和地表粗糙度具有一定的相关性外，还受气候和地形等因素影响，后续的研究需结合多种土壤水分影响因子进行反演，以提高土壤水分反演精度。

参考文献

[1] 荆雪娟. 基于Sentinel-1A SAR与Landsat-8 OLI影像的土壤含水量反演:北麓河流域[D]. 北京:中国地质大学(北京),2019

[2] 郭晓娟,周妍妍,郭建军,等. 疏勒河流域土壤含水率反演[J]. 干旱区研究,2018,35(6):1 317-1 326

[3] 李平湘,刘致曲,杨杰,等. 利用随机森林回归进行极化SAR土壤水分反演[J]. 武汉大学学报(信息科学版),2019,44(3):405-412

[4] Petropoulos G, Carlson T N, Wooster M J, et al. A Review of Ts/VI Remote Sensing Based Methods for the Retrieval of Land Surface Energy Fluxes and Soil Surface Moisture[J]. Progress in Physical Geography,2009,33(2):224-250

[5] Paloscia S, Pettinato S, Santi E, et al. Soil Moisture Mapping Using Sentinel-1 Images: Algorithm and Preliminary Validation[J]. Remote Sensing of Environment,2013,134:234-248

[6] 艾璐,孙淑怡,李书光,等. 光学与SAR遥感协同反演土壤水分研究进展[J]. 自然资源遥感,2021,33(4):10-18

[7] 王龙,宫辉力,潘云,等. 基于Sentinel-1 SAR数据的锡林浩特市典型草原土壤水分反演[J]. 干旱气候,2019,37(6):979-986

[8] 侣文娜,何敏,王茂山. 基于Sentinel-1的典型地物时间失相干分析[J]. 地理空间信息,2022,20(2):22-26

[9] 郑少兰,刘龙威. 单时相Sentinel-1A卫星SAR数据水稻信息提取[J]. 地理空间信息,2020,18(4):61-64

[10] 童李霞,燕琴,骆成凤,等. 基于NDWI分割与面向对象的水体信息提取[J]. 地理空间信息,2017,15(5):57-59

[11] Attema E P W, Ulaby F T. Vegetation Modeled as a Water Cloud[J]. Radio Science,1978,13(2):357-364

[12] Jackson T J, CHEN D Y, Cosh M, et al. Vegetation Water Content Mapping Using Landsat Data Derived Normalized Difference Water Index for Corn and Soybeans[J]. Remote Sensing of Environment,2004,97(4):475-482

[13] Bindlish R, Barros A P. Parameterization of Vegetation Backscatter in Radar-based, Soil Moisture Estimation[J]. Remote Sensing of Environment,2001,76(1):130-137

[14] 苏日图. 融合多源数据的草原区土壤水分监测方法研究[D]. 呼和浩特:内蒙古工业大学,2018

[15] Fung A K, LI Z, CHEN K S. Backscattering from a Randomly Rough Dielectric Surface[J]. IEEE Transactions on Geoscience and Remote Sensing,1992,30(2):356-369

[16] Fung A K, CHEN K S. An Update on the IEM Surface Backscattering Model[J]. IEEE Geoscience and Remote Sensing Letters,2004,1(2):75-77

[17] Baghdadi N, Holab N, Zribi M. Calibration of the Integral Equation Model for SAR Data in C-band and HH and VV Polarizations[J]. International Journal of Remote Sensing, 2006,27(4):805-816

[18] HUANG G, HUANG G B, SONG S, et al. Trends in Extreme Learning Machines: A Review[J]. Neural Networks, 2015,61:32-48