

基于风云气象卫星的土壤湿度数据降尺度方法研究

盛佳慧^{1,2,3}, 饶鹏^{1,2*}

- (1. 中国科学院上海技术物理研究所, 上海 200083;
2. 中国科学院智能红外感知重点实验室, 上海 200083;
3. 中国科学院大学, 北京 100049)

摘要: 针对被动微波土壤湿度数据空间分辨率较低的问题, 分别基于随机森林、多项式拟合及DISPATCH等统计学和物理模型, 融合可见光、热红外和地表高程参量对风云三号B星(FY3B)微波土壤湿度数据进行降尺度, 使其空间分辨率从25 km提高至1 km。同时, 考虑FY3B、与相关输入数据源过境时间不匹配现象, 设置升降轨共计四组对照实验, 对比分析FY3B降尺度的最优化数据组合。采用2015年4月1日至2016年12月31日的REMEDHUS土壤湿度原位站点及ECA&D气象站点数据验证, 结果显示随机森林方法综合降尺度精度最高, 模型拟合效果最好。此外, 采用FY3B升轨数据降尺度效果更优。

关键词: 土壤湿度; FY3B/MWRI; MODIS; 降尺度; 随机森林; DISPATCH; 多项式拟合; REMEDHUS
中图分类号: TP79; S152.7 **文献标识码:** A

The research on downscaling methods based on Fengyun meteorological satellite soil moisture data

SHENG Jia-Hui^{1,2,3}, RAO Peng^{1,2*}

- (1. Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;
2. Key Laboratory of Intelligent Infrared Perception, Chinese Academy of Sciences, Shanghai 200083, China;
3. University of the Chinese Academy of Sciences, Beijing 100049, China)

Abstract: In view of the low spatial resolution of passive microwave soil moisture (SM) data, statistical and physical models including random forest (RF), polynomial fitting and DISPATCH are utilized to disaggregate the FY3B microwave SM product from 25 km to 1 km with the synergistic application of Optical/Thermal infrared (TIR) observations and surface elevation parameters. Meanwhile, considering different overpass times of FY3B and other relevant input data source observations, four data combinations are separately used to derive the spatially downscaled SM with above three downscaling method, and the optimized data combination of FY-3B downscaling is proposed by comparison and analysis. Validation is performed from April 1, 2015 to December 31, 2016 with the in-situ measurements of REMEDHUS network and the precipitation time series of ECA&D meteorological site. Experimental results show that RF-based method can achieve the highest comprehensive downscaling accuracy and the best model fitting effect. In addition, the effect of applying FY-3B ascending data to downscale turns out to be better.

Key words: Soil moisture, FY3B/MWRI, MODIS, downscaling, random forest, DISPATCH, polynomial fitting, REMEDHUS

PACS.: 07.07.Df, 43.60.Rw

收稿日期: 2019-11-09, 修回日期: 2020-07-10

Received date: 2019-11-09, Revised date: 2020-07-10

基金项目: 中国科学院先导培育计划项目(09KCE043N2)

Foundation items: Supported by Pilot Cultivation Program of the Chinese Academy of Sciences (09KCE043N2)

作者简介(Biography): 盛佳慧(1996-), 女(满族), 黑龙江绥化人, 硕士研究生, 现从事遥感图像处理技术研究。E-mail: shengjiahui@mail.sitp.ac.cn

*通讯作者(Corresponding author): E-mail: peng_rao@mail.sitp.ac.cn

引言

土壤湿度(SM)表征包含在土壤颗粒间隙的水分含量^[1]。地表土壤水分是调节地表过程与大气之间联系的重要参量,也是控制地气水分与热量交换过程的关键要素。地表土壤水分时空动态变化与地球气候、生态、生物化学和农业环境息息相关。SM观测在全球的气候变化预测、数值天气模拟、农业干旱预警、作物检测估产以及洪水灾害监测等应用领域具有重要意义^[2]。SM观测需要具备较高的时空分辨率、较大的空间覆盖度和时间跨度,以满足大范围与高时相的监测需求。目前,获取土壤表面湿度主要通过原位传感器站点实测、机载传感器以及卫星观测三种方法^[3]。原位土壤湿度传感器网络具有较高的测量精度和时间采样,然而站点分布稀疏不均、空间分辨率低、成本过高、工作效率低等缺点严重制约着传感器网络测量方法的应用^[4]。机载传感器主要包括被动微波辐射计和散射计,能够提供较高的时空分辨率和精度,但由于受到飞行条件的限制,无法实现观测覆盖全球范围^[5]。近年来,通过遥感卫星获取SM卫星的数据逐渐成为趋势,SM卫星能够具备适中的重访周期、覆盖全球范围的观测能力以及由中到高的反演精度。如今,许多卫星可以提供接近实时的SM产品,包括主被动观测结合的SMAP卫星、土壤水分海洋盐度SMOS卫星、Sentinel-1卫星等。与光学遥感相比,微波对土壤水分更加敏感,且微波具有一定的穿透能力,不受云、雨、雾的限制,微波遥感在时空尺度上反演的SM数据更加准确和连续。然而,微波遥感数据空间分辨率较为粗糙,难以满足精细农业等应用需求。其中,近年来新兴的风云三号卫星SM产品FY3B空间分辨率较低,极大地限制了该产品在土壤湿度研究方面的应用^[6]。

近年来,国内外学者采用各种降尺度方法提高SM产品的空间分辨率。当前主流的降尺度方法可以分为基于卫星的方法^[7],结合地理信息数据的方法和基于模型的方法。基于卫星的方法主要包含主被动微波融合和可见光热红外与微波融合的方法。为了充分利用辐射计和雷达观测数据的优点,研究学者和分别提出变化检测算法^[8]和贝叶斯归并算法^[9],协同使用辐射计和雷达数据进而获得高空间分辨率的SM数据^[10]。应用变化检测算法使用L波段辐射计和雷达观测在SMEX02区域进行降尺度试验^[11]。通过OSSE观测系统进行模拟观测,研究

表明主被动结合的降尺度算法能够较好地表达SM的空间分布细节^[12]。针对被动微波土壤湿度产品对降尺度结果的影响,改进的基线方法通过雷达后向数据对亮度温度数据进行降尺度处理^[13],从而获得降尺度后的土壤湿度数据^[14]。另外有研究^[15]直接通过雷达土壤湿度产品对散射计土壤湿度产品进行降尺度,研究结果表明基于主被动融合的土壤湿度数据与简化的基于小波的图像增强方法有关^[16]。相比于微波遥感,可见光、热红外遥感具有能够提供高分辨率陆地表面参数的优势。大量的研究尝试基于可见光或热红外获得的植被覆盖和陆地表面温度信息,基于这两种辅助信息获得降尺度的微波遥感的土壤湿度数据^[17]。这些方法的核心思想是从高分辨率的可见光热红外数据中获得降尺度因子,通过降尺度因子改善低分辨率的微波土壤湿度数据空间差异的表达。基于表面温度和植被指数的三角特征空间^[18],另外有研究开发了基于经验的多项式拟合的降尺度方法^[19,20],这种方法通过低空间分辨率的陆地表面温度、植被指数和表面反照率数据构建多项式函数,再通过高分辨率陆地表面温度、植被指数和表面反照率数据得出降尺度后的土壤湿度,这种方法之前应用于SMOS、AMSR-E土壤湿度产品获得了较好的降尺度效果。还有学者开发了基于热红外光谱的降尺度方法(DISPATCH)^[21],该方法主要基于植被蒸发效应关联高空间分辨率的土壤湿度数据与粗糙的土壤湿度数据。DISPATCH方法以土壤蒸发过程为基础,将光学数据和近地表土壤湿度数据关联起来,是一种基于物理的方法。有研究基于人工智能技术结合MODIS表面温度数据对SMOS土壤湿度产品进行降尺度研究^[22],主要采用的人工智能技术包括人工神经网络^[23]、支持向量机和相关向量机^[24]。机器学习方法中,随机森林方法已在建筑物变形、滑坡监测等非线性事件处理中得到广泛应用。然而在土壤湿度降尺度方法研究较少采用随机森林方法,尤其在风云三号气象卫星FY3B土壤湿度产品降尺度方面^[25]。

基于上述研究,本文基于DISPATCH、随机森林以及多项式拟合三种方法对风云三号FY3B土壤湿度产品进行降尺度方法研究,将原始产品空间分辨率从25 km提高至1 km。DISPATCH与多项式拟合方法在土壤湿度降尺度方面较少应用在FY3B数据产品上。近年来,基于机器学习方法进行土壤湿度

降尺度成为一种新型的降尺度方法^[26]。其中,随机森林方法具有能够较好模拟变量间复杂的非线性关系的优势,从而被广泛应用于分类与回归任务^[27]。本文对比三种方法应用于FY3B土壤湿度产品的降尺度效果,根据风云三号B星、Terra和Aqua卫星平台过境时间不匹配现象,设置升降轨共计四组对照实验,对比分析FY-3B降尺度的最优化数据组合。其中,数据组合分别为AM+Terra、AM+Aqua、PM+Terra、PM+Aqua。为验证DISPATCH、随机森林以及多项式拟合方法降尺度的结果,将REMEDHUS原位站点网络和欧洲气候评估数据作为参考数据,对三种方法的降尺度结果进行时空精度验证。

1 数据和研究区域

1.1 研究区域及验证数据

考虑到伊比利亚半岛通常为晴朗的天空条件,本次实验研究区域选取了几乎整个伊比利亚半岛。整个研究区域如图1所示。根据MODIS的土地覆盖数据,本文研究区域的地表覆盖类型是混合的,但该地区的主要土地覆盖类型为农田、开阔灌木和木质稀树草原,部分森林位于中部山区。整个研究区域海拔为9 m至2 568 m,平均海拔约729 m。气候类型属于半干旱的地中海大陆性气候,夏热冬凉。

本研究区域内的密集土壤水分观测网(REMEDHUS)位于西班牙杜罗河盆地中部半干旱

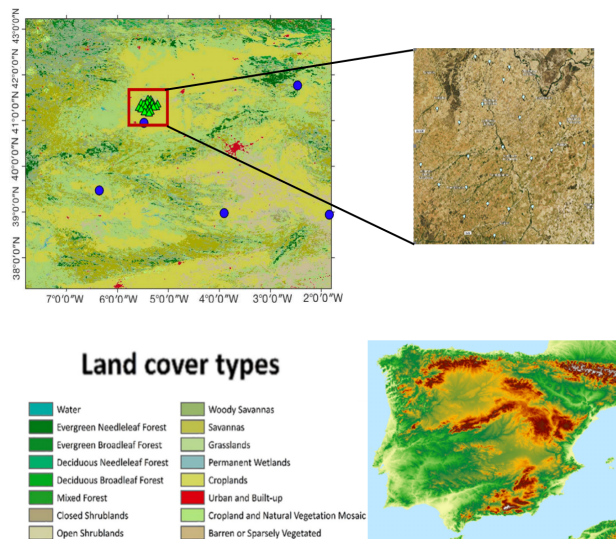


图1 研究区域及高程信息、REMEDHUS原位站点网络和气象站的分布情况

Fig. 1 Study area with elevation and the distribution of the REMEDHUS network and the meteorological stations

区域。REMEDHUS网络由20个站点组成,能够提供小时为采样单位的土壤湿度数据和深度为5cm的地表温度的测量值。为从时间尺度评估降尺度结果的可靠性,本文采用欧洲气候评估数据集(ECA&D)提供的五个气象站的日降水量数据作为时间尺度验证数据。各气象站所在位置分别显示在图1中。日降水验证可以反映出降水对土壤湿度的影响和降尺度结果的时间尺度表达水平,此验证过程对于降尺度的验证非常必要。本文所有数据研究时间区间选取2015年4月1日至2016年12月31日,包括REMEDHUS土壤湿度原位站点及ECA&D气象站点数据。

1.2 FY3B土壤湿度产品

风云三号B星是中国第二代极地轨道气象卫星,由中国国家航天局于2010年11月5日发射。卫星平台备有被动式微波辐射计(MWRI),MWRI传感器主要任务是通过接收来自地球表面和大气的水平、垂直两个极化的电磁辐射信息^[15],反演出全球降水、云和大气中水汽含量、地表植被、海表温度及土壤湿度等信息。传感器探测频率由10.7 GHz至89.0 GHz不等。轨道高度约为836km,重访周期约为3天2次。风云三号气象卫星每天两次扫过地面同一地点,即降轨和升轨。卫星降轨和升轨过境时间分别为上午1时40分和下午1时40分。本文用PM表示FY3B升轨数据,AM表示FY3B降轨数据。本研究采用FY3B的2L级产品,土壤湿度(VSM)产品分辨率为25 km×25 km,选取FY3B数据时间范围为2015年4月1日至2016年12月31日。FY3B土壤湿度产品可在中国国家卫星气象中心网站免费下载。

1.3 MODIS产品

Aqua和Terra是地球观测系统中两颗重要的卫星。其中,Terra为上午星,过境时间为地方时上午10:30,Aqua为下午星,过境时间为地方时下午1:30,两颗卫星均载有中分辨率成像仪(MODIS)。根据大量降尺度研究表明,土壤湿度与地表温度(LST)、增强型植被指数(EVI)、叶面积指数(ALB)、归一化植被指数(NDVI)、归一化水指数(NDWI)有关。本文主要采用Terra与Aqua卫星的MODIS观测数据获得研究所需的陆地表面变量。LST数据来源于MOD11A1和MYD11A1产品,空间分辨率为1 km,时间分辨率为一天。EVI和NDVI数据获得自Terra卫星的MOD13A2产品,空间分辨率为1 km,

时间分辨率为16天。本研究将空间分辨率为500 m的MCD43A3产品重采样至1 km获得反照率ALB数据,时间分辨率为16天。空间分辨率为500 m,时间分辨率为8天的地表反射率产品MOD09A1经过计算得到NDWI数据。从MCD15A2产品获得LAI参数,空间分辨率为1 km,时间分辨率为8天。所有数据都可以从美国陆地进程分布式活动存档中心(LP DAAC)下载。在之前敏感性分析研究的基础上,考虑到白天的LST对土壤湿度敏感度较高,本实验只选取Terra和Aqua卫星白天的LST数据应用于本文中。由于MODIS和FY3B卫星平台不同,考虑升降轨过境时间的差异,本文假设MODIS和FY3B的升降轨过境时间在土壤湿度分布空间上是持续性的。

1.4 地形数据

数字高程模型(DEM)数据采集于航天飞机雷达地形任务(SRTM)中,该任务能够提供空间分辨率分别为30 m、90 m和1 km的数据集。考虑到FY3B土壤湿度产品的原始分辨率,本次实验选择版本号为2.1的DEM数据版本,分辨率为30 arcsec(1 km)。为清晰展示本次研究采用的数据集,表1列出本研究所涉及的全部数据集的名称、空间分辨率、变量描述以及产品瓦片数。值得注意的是,表中1 km的NDWI数据是基于分辨率为500 m的MOD09A1进行空间重采样计算出来的。

表1 实验数据集概述

Table 1 Overview of the datasets used in this study

数据集	空间分辨率	描述	瓦片数
FY3B	25 km	土壤湿度 (SM)	1945
MOD11A1	1 km	陆表温度 (LST)	3982
MYD11A1	1 km	陆表温度 (LST)	4000
MOD13A2	1 km	归一化植被指数 (NDVI) 增强型植被指数 (EVI)	252
MCD43A3	500 m	地表反照率 (ALB)	4032
MOD09A1	500 m	归一化水体指数 (NDWI)	504
MCD15A2	1 km	叶面积指数 (LAI)	504
SRTM	1 km	地形高程	2

2 降尺度方法

2.1 基于随机森林的降尺度原理

将FY3B微波土壤湿度与可见光、热红外和地表高程进行融合,依据的是其物理意义和统计学方法。在物理意义上,可见光波段反映了地表植被覆盖情况,而植被通过影响太阳入射辐射、土壤蒸散

发进而影响了土壤湿度。热红外反映了地表温度情况,温度越高,土壤水蒸散发效应越显著,地表越干旱。地表高程则通过影响地表温度从而反映到土壤湿度上。在统计学方法上,本文利用随机森林模型研究构建土壤湿度与上述参量的非线性关系。由于上述参量空间分辨率较FY3B微波土壤湿度产品更高,因此通过上述融合,可以将FY3B土壤湿度产品分辨率提高,从而达到降尺度的效果,如图2所示。

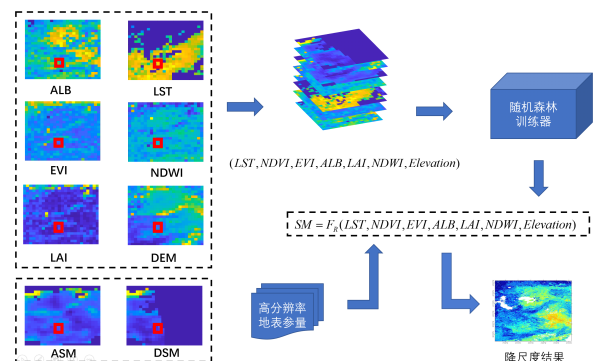


图2 可见光、红外、地表参量融合流程图

Fig. 2 Fusion flowchart of Optical/Thermal infrared (TIR) observations and surface elevation parameters

融合可见光、热红外和地表高程参量进行土壤湿度降尺度研究核心思想是获得地表温度、太阳辐射、植被覆盖及海拔高程等地表参数,结合各参数的降尺度因子提高微波土壤湿度的空间分辨率。融合可见光、热红外和地表高程参量的具体流程为:先通过土壤湿度原始数据与可见光、热红外和地表高程等参量构建关系表达式,以确定回归系数,即降尺度因子。通过应用具有降尺度因子的关系表达式获得高分辨率土壤湿度。本文利用机器学习方法中的随机森林算法将影响土壤湿度的物理参量通过黑盒模型构建其与土壤湿度的联系,从而提升遥感土壤湿度产品的空间分辨率。

随机森林(RF)是一种基于多决策树的集成学习算法。具体步骤是将训练数据划分为许多回归树,其中每棵树由包含三分之二训练数据的bootstrap样例构建。RF方法的另一个特点是bootstrap采样过程中排除的剩余样本将作为Out-Of-Bag样本(OOB)。OOB可用于验证所构建模型的性能,相当于机器学习中的k倍交叉验证,不需要额外的计算成本。基于RF方法的FY3B土壤湿度降尺度方法的研究建立在已知前提下:(1)机器学习算法在

各种基于遥感卫星的土壤湿度数据产品降尺度方法中得到了广泛的应用,且取得了较好的效果。(2)土壤湿度变化是一个复杂的多因子交互作用的土壤水文过程。已有研究表明EVI、LAI、ALB、NDVI、DEM和NDWI等参数均对土壤持水能力有一定影响。(3)空间分辨率为25 km的FY3B土壤湿度产品单像素值代表像素内部的整体属性,而不是区域内的极端情况。在RF参数选择中,有一些重要的参数,包括树的数量,以及每次分割时选择的变量的数量。本文根据OOB误差的稳定特性,采用网格搜索的方法训练大量数据训练模型,确定算法参数设置值,网格搜索流程如图3所示。

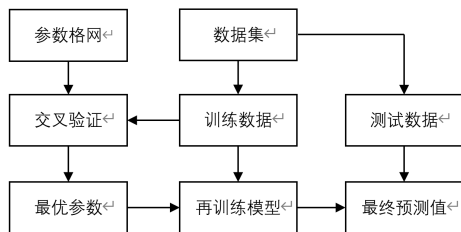


图3 网格搜索流程图

Fig. 3 Flowchart of grid search

首先,首先切分原始数据,先选定较大范围参数列表,逐步细化,创建并调用机器学习网格搜索对象,对网格搜索的实例对象进行验证数据模拟,从而得到最佳模型及参数。网格搜索能够实现自动调参,输入相关参数即可得出最优化结果和参数。细化后的参数组合拟合效果如图4所示。验证分数均值越大,方差值越小表示相应的参数组合越适合于数据集。当分割点的变量数设为7,树数目设置为1500时验证分数均值为0.437,方差值为0.134为网格搜索输出的最优参数组合。本文关键参数分割点的变量数和树的数目分别设置为7和1500。网格搜索结果由图3中红色框线表示。

采用LST、EVI、LAI、ALB、NDVI、DEM和NDWI陆地表面变量构建基于RF方法的土壤湿度关系模

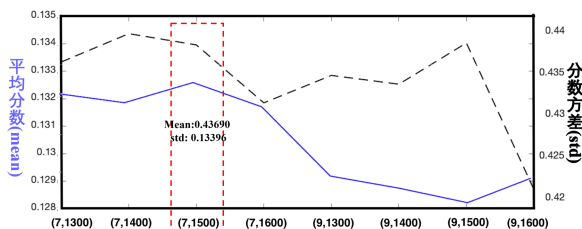


图4 网格搜索结果图

Fig. 4 The result of grid search

型。利用RF回归方法,用非线性函数(F_R)构建陆地表面变量与土壤湿度之间的复杂非线性关系如式所示。

$$SM = F_R(LST, NDVI, EVI, ALB, LAI, NDWI, Elevation), (1)$$

基于RF方法的降尺度流程中包含的数据的预处理等具体细节如下:

(1)陆地表面数据重采样。由于陆地表面变量LST、EVI、LAI、ALB、NDVI、DEM和NDWI原始数据时空分辨率存在差异,为后续将数据直接输入到机器学习算法中进行训练,故首先基于平均值采样的方法将原始数据的空间分辨率统一至1 km。陆地表面数据重采样主要采用最近邻重采样的方法。最近邻重采样方法核心思想为将目标各像素点的灰度值选取源图像中与其最邻近像素的灰度值。部分陆地表面数据重采样结果如图5所示,右上角小图为局部放大图。由于本文选取的陆地表面源数据分辨率即为1 km,故重采样前后空间分布模式变化较小。

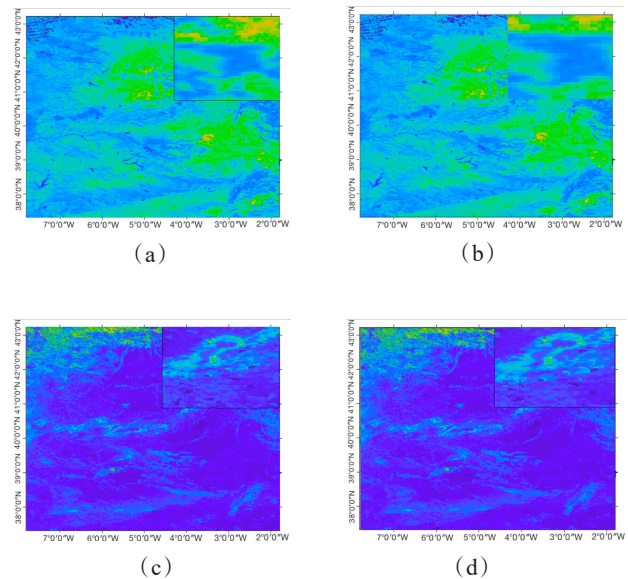


图5 陆地重采样结果对比图,(a)原始ALB数据,(b)重采样ALB数据,(c)原始NDWI数据,(d)重采样NDWI数据

Fig. 5 The results of resampling land surface data, (a) original ALB data, (b) resampling ALB data, (c) original NDWI data, (d) resampling NDWI data

(2)基于MODIS 44W空间分辨率为250 m的陆地水模数据去除研究区域中的水体区域,排除水体区域对实验造成不必要的影响。根据卫星过境时间的差异,将所有数据分为四组AM+Terra、AM+

Aqua、PM+Terra、PM+Aqua 进行对比试验,其中 AM、PM 分别表示 FY3B 数据降轨和升轨过境时间所对应的土壤湿度数据。

(3)在 25 km 的空间尺度下,基于 RF 的方法在 2015 年 4 月 1 日至 2016 年 12 月 31 日时间范围内构建基于 AM+Terra、AM+Aqua、PM+Terra、PM+Aqua 数据组的 FY3B 土壤湿度与陆表变量模型关系。

(4)基于 RF 方法训练得到的模型,对 1 km 空间分辨率的陆地表面数据进行水体去除处理,将处理结果作为训练模型的输入变量,进而计算得到降尺度后的高空间分辨率的土壤湿度数据。基于 RF 的土壤湿度降尺度流程如图 6 所示。

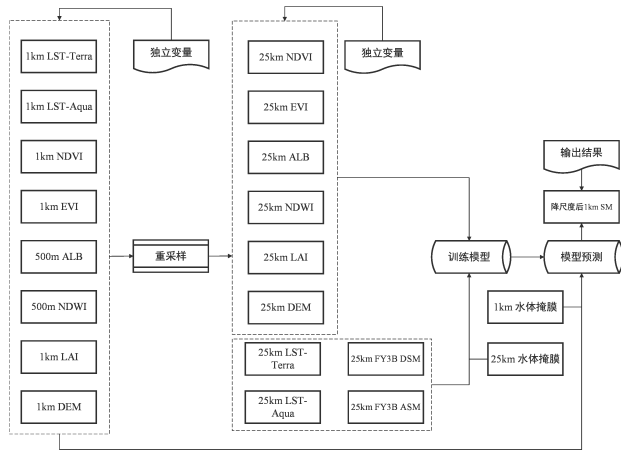


图 6 基于随机森林方法的降尺度流程图

Fig. 6 Downscaling flowchart based on random forest method

2.2 基于多项式回归的降尺度方法

大量降尺度方法尝试基于植被覆盖和陆表温度信息进行土壤湿度产品的降尺度,这些方法的核心思想是获得高空间分辨率的可见光和红外数据降尺度因子,利用降尺度因子表征高分辨率土壤湿度数据与原始粗糙的微波土壤湿度数据之间的变异值。基于广泛使用的地表温度和植被指数三角特征空间(图 7),有学者提出了一种经验多项式拟合降尺度方法^[18, 28],研究表明此方法与传统线性回归模型降尺度方法相比,基于多项式回归模型获得的降尺度结果更接近于实测值,降尺度结果更为准确。此方法基于可见光和热红外的地表温度(LST)、植被指数和地表反照率数据构建多项式函数表达高空间分辨率的土壤湿度数据。

$$SM = F_p(LST, NDVI, ALB) \\ = a_{000} + a_{100}T_s + a_{010}NDVI + a_{001}ALB + a_{200}T_s^2 \\ + a_{020}NDVI^2 + a_{002}ALB^2 + a_{110}T_s * NDVI \\ + a_{101}T_s * ALB + a_{011}ALB * NDVI \quad (2)$$

这种方法的具体流程如下:首先,将粗分辨率陆地表面数据输入到多项式模型中,确定多项式的回归系数。之后,根据计算得到的粗分辨率回归系数,将高分辨率的陆地表面数据输入到多项式表达式中,从而计算出高空间分辨率土壤湿度值。其中,多项式函数一般表示为式(2)。式中的 SM 是土壤湿度值, LST 和 NDVI 是陆地表面温度和归一化植被指数变量, a_{ijk} 表示回归系数, ALB 是经过平均采样方法处理后的地表反照率变量。

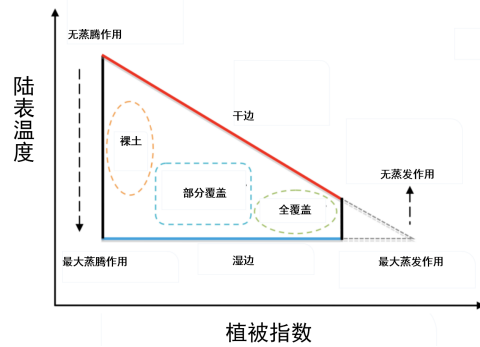


图 7 LST-fv 二维特征空间

Fig. 7 The two-dimensional feature space of LST and f_v

2.3 DISPATCH 降尺度方法

DISPATCH 算法主要原理是以 SEE 指数为基础,表征 FY3B 单像元内的土壤湿度变异表征因子。土壤水蒸散发效率(SEE)是基于热红外光谱辐射而推估高空间分辨率土壤湿度的指数,具备千米级别的空间分辨率,用来表征微波遥感土壤湿度单像元内土壤湿度的空间变异特征。SEE 指数构建的基本原理是从混合像元的地表蒸散发中剔除植被蒸散发,保留土壤成分,并建立该成分与表层土壤湿度的相关性。由于地表蒸散发由地表温度推估得到,SEE 指数利用了土壤在热红外光谱中的光谱特性。SEE 构建可遵循如式(3)所示:

$$SEE_{HR} = \frac{(T_{s,max} - T_{s,HR})}{(T_{s,max} - T_{s,min})} \quad (3)$$

式(3)中的土壤和植被的温度是根据 LST 和 NDVI 分别计算得到的。其中,土壤温度表示如式(4)所示。

$$T_{s,HR} = \frac{(T_{MODIS} - f_{v,HR}T_{v,HR})}{(1 - f_{v,HR})} \quad (4)$$

其中, T_{MODIS} 为 MOD11A1 地表温度产品提供的 LST 参数, f_v 表示植被丰度。植被丰度计算公式如式(5)所示。

$$f_{v,HR} = \frac{NDVI_{MODIS} - NDVI_s}{NDVI_v - NDVI_s} \quad (5)$$

$NDVI_{MODIS}$ 表示纯净裸土的 NDVI 值, 此处设置为 0.15。 $NDVI_s$ 为纯净植被的 NDVI 值, 并设置为 0.9。植被温度 $T_{v,HR}$ 的计算遵循“沙漏”算法, 如图 8 所示。

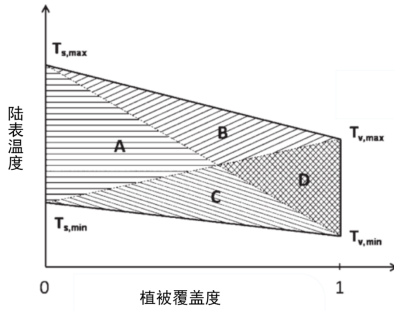


图 8 “沙漏”算法原理图

Fig. 8 Schematic diagram of "hourglass" method

图中特征空间被分成 4 个区域分别为 A、B、C 和 D。其中区域 A 表示地表温度主要由土壤水分蒸发导致; 区域 B 和 C 同时受到土壤水分蒸发和植被蒸腾作用双重胁迫, 区域 D 表示对土壤湿度没有较高的敏感性。若数据点位于区域 A, 则植被温度可由式(6)表示, 以此类推。式中的 $T_{v,min}$ 与 $T_{v,max}$ 分别为在最少和最多水分胁迫下的植被温度。

$$T_{v,HR} = (T_{v,min} + T_{v,max})/2 \quad (6)$$

$$T_{v,HR} = (T_{v,min,HR} + T_{v,max})/2 \quad (7)$$

$$T_{v,HR} = (T_{v,min} + T_{v,max,HR})/2 \quad (8)$$

$$T_{v,HR} = (T_{v,min,HR} + T_{s,max,HR})/2 \quad (9)$$

土壤与植被的最值温度计算可以遵循以下计算流程: (1) $T_{v,min} = T_{MODIS,min}$; (2) 如果 $T_{MODIS,min}$ 对应像元植被丰度较低, 则 $T_{s,min} = T_{MODIS,min}$; (3) 如果 $T_{MODIS,min}$ 对应像元的植被丰度较高, 则 $T_{s,min}$ 设为 $T_{s,HR}$ 所对应所有像元的最小值; (4) 当 $T_{MODIS,max}$ 对应像元植被丰度较低, 则 $T_{s,max} = T_{MODIS,max}$, 且 $T_{v,max}$ 设为 $T_{v,HR}$ 所对应所有像元的最大值; (5) 如果 $T_{MODIS,max}$ 对应像元的植被丰度较高, 则 $T_{v,max} = T_{MODIS,max}$, 且 $T_{s,max}$ 设为 $T_{s,HR}$ 所对应所有像元的最大值。

假设 SEE 指数与 FY3B 土壤湿度存在相关性, DISPATCH 算法是以 SEE 指数为基础, 将 25 千米空间分辨率的 FY3B 数据降尺度至 1 千米分辨率。土壤湿度的变异量可以从 SEE 域转换到土壤湿度域, 并用一阶泰勒公式展开如式(10)所示。

$$SM_H = SM + \frac{\partial SM}{\partial SEE} (SEE - SEE_{mean}) \quad (10)$$

3 结果分析

3.1 降尺度模型对比

考虑 FY-3B 与相关输入数据源过境时间不匹配现象, 设置升降轨共计四组对照实验, 分别采用三种方法进行降尺度。四组对照试验数据分别为 AM+Aqua、AM+Terra、PM+Aqua、PM+Terra。其中 AM 表示 FY3B 降轨对应的土壤湿度值, PM 表示升轨土壤湿度值。Aqua 表示 MODIS/Aqua 卫星的 LST 及其他表面变量, Terra 表示 MODIS/Terra 的 LST 及其他表面变量。基于以上四组数据计算出三种方法降尺度后的土壤湿度数据与原始数据的相关系数 (R) 与均方根误差 (RMSD) 值, R 与 RMSD 值计算公式如式(11)(12)所示。两式中的 N 均表示当天土壤湿度值数量。式(11)中, i 表示数据序号, SM_H 表示第 i 个降尺度后土壤湿度值, SM_i 表示第 i 个原始土壤湿度值, $\overline{SM_H}$ 和 \overline{SM} 分别表示 N 个降尺度后及原始土壤湿度值的均值, 表示 N 个原始土壤湿度均值, σ_H 和 σ 分别表示降尺度及原始土壤湿度标准差。式(12)中, SM_H 与 SM 分别表示降尺度后与原始土壤湿度值多项式拟合、RF、DISPATCH 方法计算结果如图 5-7 所示。

$$R = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{SM_H_i - \overline{SM_H}}{\sigma_H} \right) \left(\frac{SM_i - \overline{SM}}{\sigma} \right) \quad (11)$$

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (SM_H - SM)^2} \quad (12)$$

基于三种方法估算出的 SM 与 FY3B 原始 SM 在研究时间范围内的大部分一致性较好。然而依然存在某天具有较差的相关性。其中, 多项式拟合方法应用在研究时间范围内 R 和 RMSD 的变化如图 9 所示。由图可知, 基于 PM + Terra 和 PM + Aqua 数据组合比 AM + Terra 和 AM + Aqua 数据组合具有更好的估计 SM 的能力, 且 R 值更高。其中, PM + Terra 和 PM + Aqua 的两组 R 平均值分别为 0.794 和 0.795。AM + Terra 和 AM + Aqua 两组 R 平均值分别为 0.761 和 0.749。图 9 显示了四个组合的平均 RMSD 值, 也证实了上述结果。PM + Terra、PM + Aqua、AM + Terra、AM + Aqua 的 RMSD 均值分别为 0.056、0.063、0.057、0.065。

图 10 为基于四组数据采用 DISPATCH 方法对 FY3B 数据进行降尺度后的 R 和 RMSD 随时间变化图。通过图 6 可以直观地发现 DISPATCH 方法基于

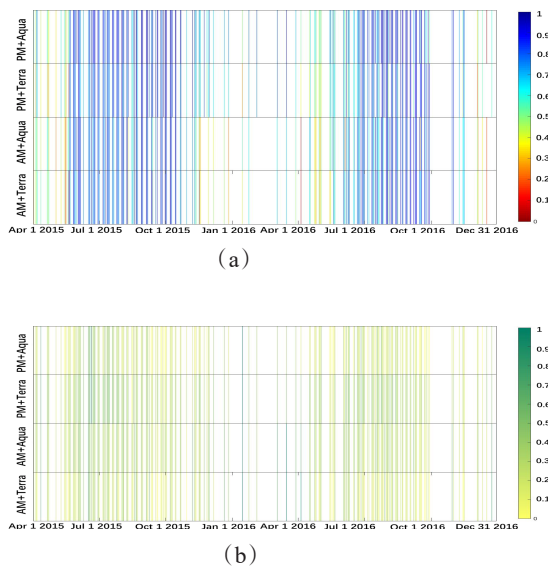


图9 基于多项式拟合方法的降尺度性能对比,(a)相关系数,(b)均方根偏差

Fig.9 The performance of polynomial-fitting method, (a)correlation coefficient, (b)RMSD

PM+Terra 和 PM+Aqua 数据组合比 AM+Aqua、AM+Terra 数据组拟合效果更好,大部分时间具有更高的 R 值。PM+Terra 和 PM+Aqua 数据组降尺度前后数据在研究时间范围内的平均 R 值分别为 0.7879 和 0.7785,AM+Terra 和 AM+Aqua 数据组在研究时间范围内的平均 R 值分别为 0.7338 和 0.7240,RMSD 值也验证了上述的结果。AM+Terra 和 AM+Aqua 数据组对应的 RMSD 比 PM+Terra 和 PM+Aqua 数据组大,按照图 6 上的数据组合从上至下的 RMSD 值分别为 0.0541、0.0540、0.0545 和 0.0541。通过与多项式方法降尺度结果进行对比,在研究范围内 DISPATCH 方法降尺度得到的土壤湿度与原始数据相关性更差。R 值较低的日子不仅仅集中在某些月份,而是更为分散地分布在研究时间范围内,R 值较低对应呈现红色竖线。相比多项式拟合方法,DISPATCH 方法的红色竖线数量明显更多。但总体来看,DISPATCH 方法对应的 FY-3B 降尺度的最优化数据组合与多项式拟合方法的结果大致相同。

由图 11 所示,基于 RF 方法降尺度前后土壤湿度数据的相关性更大。PM + Terra 和 PM + Aqua 的 R 平均值分别为 0.812 和 0.813。AM + Terra 和 AM + Aqua 的 R 平均值分别为 0.782 和 0.779。RMSD 值也有明显改善。相比多项式拟合结果,RMSD 平均值大约从 0.065 减小到 0.051。四种数据组合得

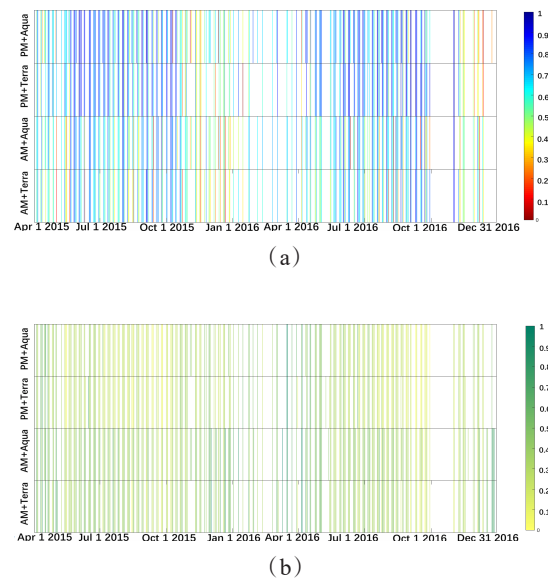


图10 基于 DISPATCH 方法的降尺度性能对比,(a)相关系数,(b)均方根偏差

Fig.10 The performance of DISPATCH method, (a)correlation coefficient, (b)The value of RMSD

到的降尺度结果的最小 R 值分别从 0.31 增加到 0.50。整体上,PM + Terra 和 PM + Aqua 数据组合的降尺度效果略好于 AM + Terra 和 AM + Aqua 组合。与此同时,采用 Terra LST 或 Aqua LST 数据的降尺度结果差异较小。总体来看,对比三种降尺度方法对应的模型拟合结果,基于 RF 方法的模型拟合效果更好,三种方法的最优数据组合呈现出相同的规律,即 PM + Terra 和 PM + Aqua 数据组合的降尺度效果略好于 AM + Terra 和 AM + Aqua 组合。不同组合之间的性能差异可以部分解释为 SM 对 LST 的敏感性不同。在白天,太阳光照引起了 LST 的变化并与 SM 高度相关。地表温度和水分剖面在清晨或夜间的平衡站趋于均匀,故白天土壤湿度对 LST 的敏感性较低。

3.2 基于随机森林方法的降尺度结果

根据 4.1 节,基于多变量拟合、DISPATCH、RF 方法构建 SM 和其他陆地表面变量之间的关系,研究结果显示 RF 方法具有最好的模型拟合效果。为了进一步评估降尺度方法的效果,将高空间分辨率陆地表面变量作为 RF 模型的输入,以获得 FY3B 的高空间分辨率土壤湿度数据。探究并对比四组数据对 RF 方法降尺度效果的影响,采样对比时间选择了 2015 年的第 136 天与第 316 天。图 12(a)和(b)显示出 2015 年第 136 天的 AM+Terra 和 AM+Aqua 数

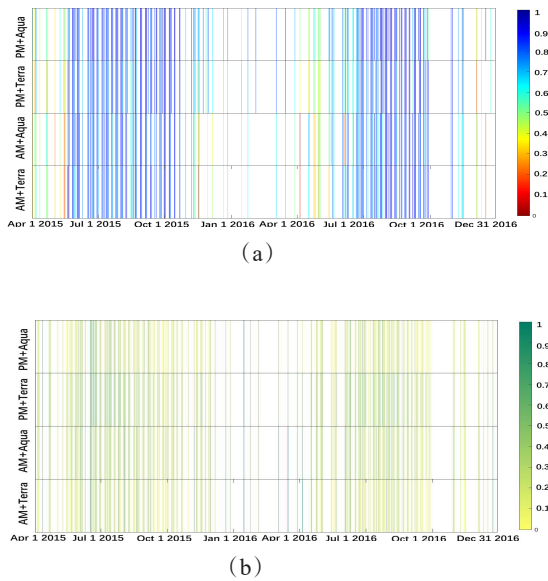


图 11 基于随机森林方法的降尺度性能对比, (a) 相关系数, (b) 均方根偏差
 Fig.11 The performance of random forest method, (a) correlation coefficient, (b) The value of RMSD

据组降尺度后土壤湿度的空间分布, 相应的原始土壤湿度数据由图 12(c) 表示。FY3B 降尺度前后 SM 值都不能完全覆盖研究区域, 这主要是由于风云三号卫星的扫描条带间隙以及 MODIS 观测部分区域出现云覆盖现象。因此, RF 模型应用于 MODIS 和 FY3B 数据均可用的像素点。降尺度 SM 结果能够有效填充由于风云卫星观测出现的扫描间隙, 需要满足的条件是在 FY3B 空白区域内, MODIS 传感器的观测是有效的。图 12(a) 和 (b) 的空白区像素数明显不同, 主要原因在于 Terra 和 Aqua 卫星的 MODIS 产品有效像素分布不同。总体上看, 虽然空白区域面积存在差异, 但是降尺度 SM 值空间分布几乎一致。

图 13 为 2015 年第 316 天针对 PM + Terra 和 PM + Aqua 数据的 RF 降尺度结果。尽管降尺度结果存在少量空白间隙, 两组数据的降尺度结果依然在 整个研究区域内呈现出十分相似的空间分布, 同时保持着与原始风云土壤湿度数据具有较好的一致性。通过 以上分析结果可以发现基于 RF 的降尺度方法能够 得到高空间分辨率的土壤湿度空间分布图。同时, 针对不同的数据组能够得到一致的空间分布。由于 MWRI 和 MODIS 传感器不同的观测系统, 如果 MODIS 观测数据是可用的, 部分 FY3B 的扫描间隙可以被部分填充。然而 MODIS 观测的云层污

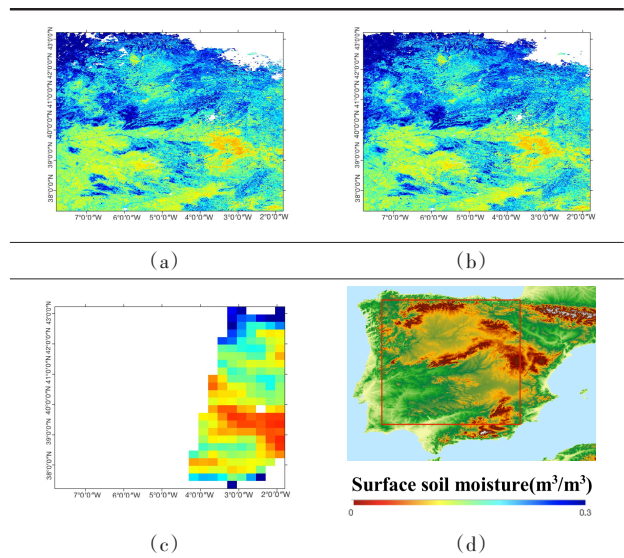


图 12 基于随机森林方法 AM+Aqua 和 AM+Terra 降尺度结果, (a) AM+Aqua, (b) AM+Terra, (c) FY3B AM
 Fig.12 Downscaling results of AM+Aqua 和 AM+Terra combinations based on random forest method, (a) AM+Aqua, (b) AM+Terra, (c) FY3B AM

染现象导致基于单一数据组降尺度结果出现大量无效值。

值得注意的是, 基于 Terra 与 Aqua 的降尺度结

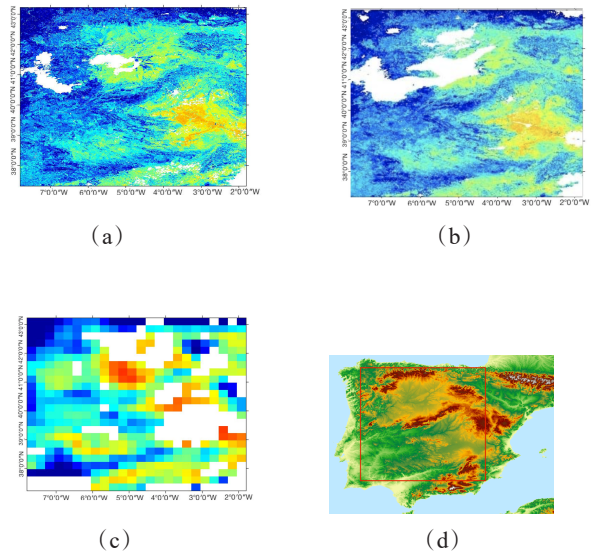


图 13 基于随机森林方法 PM+Aqua 和 PM+Terra 降尺度结果, (a) PM+Aqua, (b) PM+Terra, (c) FY3B PM, (d) 研究区域地形图
 Fig.13 Downscaling results of PM+Aqua 和 PM+Terra combinations based on random forest method, (a) PM+Aqua, (b) PM+Terra, (c) FY3B PM, (d) DEM of study area

果差异较小。针对单一数据组降尺度结果存在大量无效值现象,本文采用合成 Terra 和 Aqua 数据,采用三种降尺度方法。其中 DISPATCH 方法的四个特征区域如图 14 所示,与地表覆盖类型对比可知,特征空间中四块区域与陆地表面类型具有相关性,两者空间分布具有较大的一致性。其中,紫色区域为特征空间中的区域 A,表示地表温度主要由土壤水分蒸发导致;蓝色区域表示区域 B,绿色区域表示区域 C,区域 B 和 C 表示同时受到土壤水分蒸发和植被蒸腾作用双重胁迫;黄色区域为区域 D,表示对土壤湿度没有较高的敏感性。

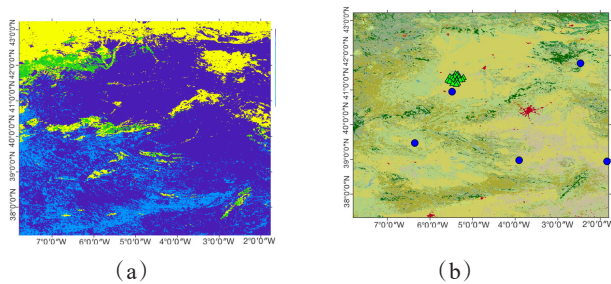


图 14 特征空间分布图, (a)特征空间分布图, (b)土地覆盖类型图

Fig. 14 Feature space distribution map, (a)Composite map based on random forest method, (b) Composite map based on polynomial-fitting method

三种方法的降尺度结果如图 15 所示,缓解了单个数据组中的云覆盖的影响。这种合成的方法主要目的在于用其他数据组的有效值填充单个数据的空白区。对比三种方法基于合成数据的降尺度结果发现,多项式拟合降尺度结果整体高于原始 SM 值。DISPATCH 方法推估出的土壤湿度空间分布图呈现出较为明显的板块效应,这主要归结于 DISPATCH 模型建立只基于 LST、NDVI 和 ALB 三个陆地表面变量,对于空间土壤湿度的变化并不能精细地捕捉到,每个像素的 SM 值极大地依赖于所在的 25 千米的原始数据值。基于 RF 的降尺度结果与原始 SM 结果具有较好的空间一致性,同时能够得到更为精细地反映出 SM 空间变化。

3.3 土壤湿度结果验证

为验证 SM 降尺度结果的精度,并进一步评估三种方法降尺度性能,本文首先基于 REMEDHUS 土壤湿度原位站点数据分别对 AM+Aqua、AM+Terra、PM+Aqua、PM+Terra 采用 RF 方法的降尺度结果

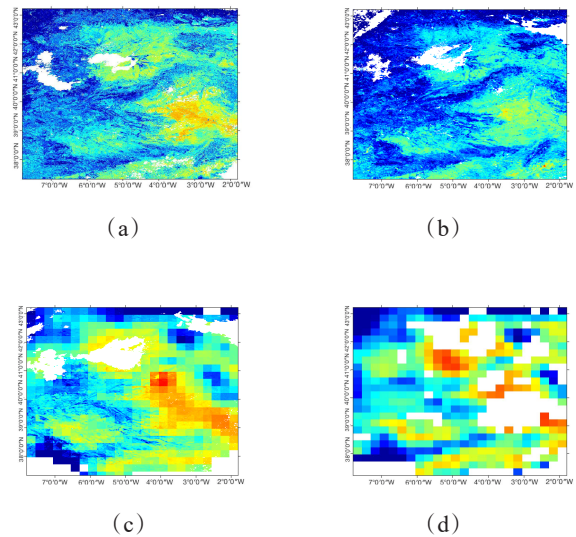


图 15 基于随机森林方法 PM+Aqua 和 PM+Terra 降尺度结果, (a)随机森林降尺度结果, (b)多项式拟合降尺度结果, (c) DISPATCH 降尺度结果, (d)原始数据

Fig. 15 Downscaling results of PM+Aqua 和 PM+Terra combinations based on random forest method, (a)Composite map based on random forest method, (b)Composite map based on polynomial-fitting method, (c)Composite map based on DISPATCH method, (d) Original data map

进行评估,对比分析 FY3B 降尺度的最优化数据组合。之后,基于整个原位站点网络数据对 DISPATCH、RF、多项式拟合三种方法降尺度结果进行单一数据组精度验证。

首先,基于原地站点网络实测数据分析基于不同数据组降尺度结果的差异。限于篇幅,只展绘了基于测站 Las Arenas 的散点图,散点图中每点对应着单日粗糙分辨率和高分辨率 SM 均有效,如图 16 和 17 所示。图 16(a)和(c)分别为原位站点实测数据与 AM+Aqua 和 AM+Terra 原始 SM 数据的散点图,图 16(b)和(d)分别为原位站点实测数据与基于 RF 方法 AM+Aqua 和 AM+Terra 降尺度数据的散点图。其中,AM+Aqua 降尺度前后 R 值从 0.68 提高至 0.76,整体增加了 0.08。RMSD 几乎保持不变,而 AM+Terra 降尺度前后 R 值显著提高,从 0.58 变化至 0.79,整体增加了 0.21。RMSD 值稍微减小,从 0.076 变化至 0.072。图 17(a)和(c)分别为 REMEDHUS 网络 Las Arenas 的原位站点实测数据与 PM+Aqua 和 PM+Terra 原始 SM 数据的散点图,图 17(b)和(d)分别为 REMEDHUS 网络 Las Arenas 的实测数据与基于 RF 方法 PM+Aqua 和 PM+Terra 降尺度

数据的散点图。其中,PM+Aqua数据组降尺度前后R值从0.59提高至0.73,增大了0.14。RMSD值降尺度前后显著减小,从0.075减小至0.064,RMSD值降尺度后降低了0.011。而PM+Terra数据组降尺度前后R值显著提高,从0.58增加至0.79,R值整体变化了0.21。RMSD值降尺度前后显著减小,从0.076减小至0.063,RMSD值整体减小了0.013。四组数据基于RF方法降尺度后R值均有上升,RMSD值呈现下降趋势。尽管图16(a)和(c)呈现出原位站点数据与FY3B AM数据相似的对比结果,由于MODIS的LST产品来自于不同的卫星(Terra星和Aqua星),可用于验证的天数稍有不同。17(a)和(c)同理。

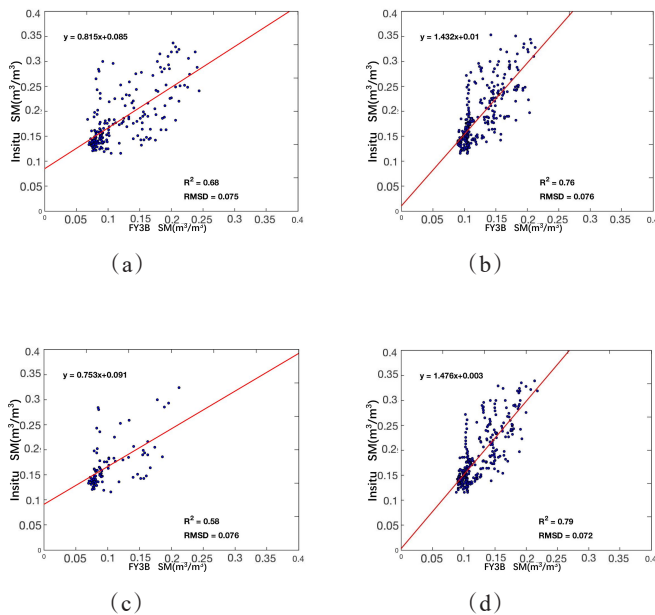


图16基于随机森林方法PM+Aqua和PM+Terra降尺度结果,(a) AM + Aqua, (b) AM + Aqua, (c) AM + Terra, (d) AM + Terra

Fig. 16 Downscaling results of PM+Aqua and PM+Terra combinations based on random forest method, (a) AM + Aqua, (b) AM + Aqua, (c) AM + Terra, (d) AM + Terra

总体来看,在AM+Aqua、AM+Terra、PM+Aqua、PM+Terra四组数据中,基于PM+Terra数据进行RF降尺度的效果最好,精度提高最为明显,R值提高最为显著,RMSD值降低最为突出。此外,在验证过程中根据对散点图进行拟合可以得出回归线的斜率和截距,通过截距值可以发现原始数据存在系统性的干燥偏倚。在进行降尺度后,这种低估效应明显

减小,这主要归因于RF方法构建的非线性多变量模型能够较好地表达SM与陆地表面变量的关系。

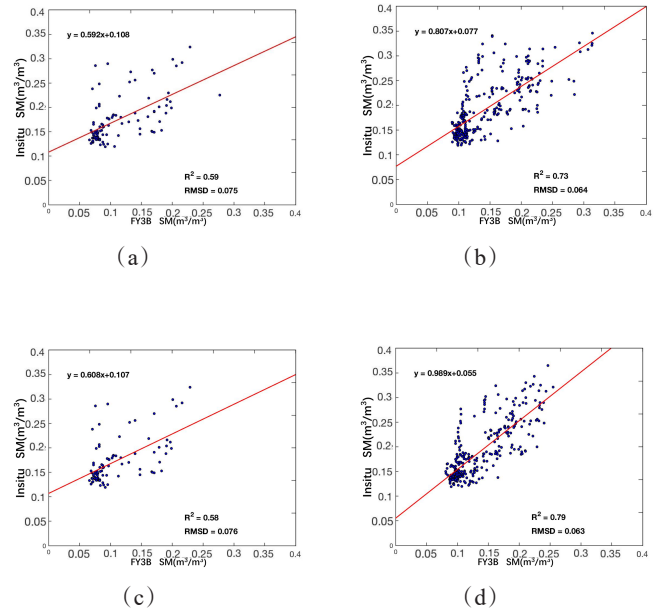


图17基于随机森林方法PM+Aqua和PM+Terra降尺度结果,(a) PM + Aqua, (b) PM + Aqua, (c) PM + Terra, (d) PM + Terra
Fig. 17 Downscaling results of PM+Aqua and PM+Terra combinations based on random forest method, (a) PM + Aqua, (b) PM + Aqua, (c) PM + Terra, (d) PM + Terra

采用整个原位SM站点网络对多变量拟合、RF和DISPATCH方法的降尺度结果进行精度验证。评估降尺度方法性能的经典指标包含均方根误差(RMSD)、相关性(R)和数据数量(Num)。此外,评估降尺度结果表征FY3B单像素土壤水分空间变化的能力是十分重要的。因此,本文除了采用常规的统计学评价指标,还采用指标bVariance评估降尺度后单像素内SM的空间代表性。bVariance表达式如式(13)所示。

$$bVariance = (STD_d - STD_{in-situ}) \times 100, \quad (13)$$

其中, STD_d 表示降尺度SM数据的标准差, $STD_{in-situ}$ 表示原地站点SM数据的标准差。bVariance值越小,表示降尺度SM值越接近原位站点数据,也意味着能降尺度SM值能够更好地表达SM的空间异质性。本研究基于R值、RMSD值、bVariance值和Num指标,采用REMEDHUS网络所有站点实测数据对三种方法降尺度结果进行精度验证,选择2015年4月1日至2016年12月31日作为验证数据时间范

围,验证结果如表2所示。

验证结果表明 DISPATCH、RF、多项式拟合三种方法的降尺度结果与原位站点实测数据均有较好的一致性。综合三种指标分析结果表明基于 RF 方法的降尺度结果更为精确,是三种降尺度方法里面最为可靠的方法,能够得到较好的降尺度效果。对比三种方法的 R 值,RF 方法与多项式方法 R 值相对较高, R 平均值分别为 0.58 与 0.63,两种方法的 R 最小值分别为 0.27 与 0.08,可以发现虽然多项式方法的平均 R 值稍大,但极值现象严重,对于某些站点应用多项式方法降尺度效果较差,主要原因归因于在地形变换较为起伏的地区不能较好地模拟 SM 与陆地表面变量之间的关系。然而,DISPATCH 方法的 R 值仅为 0.55,DISPATCH 方法的降尺度结果是三种方法中与实测数据相关性最差的。三种方法的 bVariance 值相似,表明三种方法均能较好地表现 SM 的空间变异性,三种方法的 bVariance 平均值分别为 0.022、0.023 和 0.023。然而,三种方法对应的 RMSD 指标呈现出较大的差异。与另外两种方法相比,RF 方法对应的平均 RMSD 值为 0.074,多项式拟合方法对应的 RMSD 值次之,平均 RMSD 值为

0.087,DISPATCH 方法对应的 RMSD 值最次,平均 RMSD 值为 0.098。

基于以上验证结果表明基于 RF 方法的降尺度数据在整个验证时间和大部分原位站点都有较好的验证结果,降尺度结果可能受限于原始 FY3B SM 产品的精度。除此之外,在大部分实测站点 bVariance 值越低就代表 FY3B 格网空间代表性越好,不同于 RMSE 和 R 值需要基于大量数据以具有统计意义。

3.4 基于降水数据的时间尺度验证

结合研究时间范围内的气象站点数据,对多变量拟合、RF 和 DISPATCH 方法降尺度结果进行时间尺度验证。图 18 显示出降尺度数据、原始 SM 数据与本文所选的五个气象站点的日常降雨数据随时间变化的情况。从时间尺度上进行对比,三种方法得出的降尺度数据虽然与 FY3B 原始数据之间具有一定差异,但是都能够较好地表现出时间一致性。这说明 SM 降尺度数据极大地保持了 FY3B SM 数据的时间变化特征。

针对日常降水事件,图 18 显示 SM 降尺度数据时间变化能够较好地时间尺度上表征出来。对

表2 原位数据精度验证 DISPATCH、RF、PF 降尺度结果

Table 2 In-situ validation of the downscaled result of DISPATCH、RF、PF

原位站点名称	DISPATCH				Random forest				Polynomial-fitting			
	R	bVar	Num	RMSD	R	bVar	Num	RMSD	R	bVar	Num	RMSD
Canizal	0.17	0.034	181	0.21	0.28	0.046	251	0.185	0.29	0.047	264	0.159
Carretoro	0.58	0.002	197	0.06	0.74	0.001	293	0.085	0.68	0.004	306	0.118
CasaPeriles	0.61	0.015	211	0.06	0.60	0.019	306	0.046	0.71	0.022	320	0.062
ConcejodelMonte	0.48	0.031	183	0.12	0.61	0.027	267	0.052	0.67	0.027	279	0.044
ElCoto	0.53	0.018	205	0.06	0.62	0.002	296	0.085	0.48	0.006	308	0.107
ElTomillar	0.50	0.018	198	0.07	0.49	0.003	272	0.095	0.68	0.004	284	0.128
Granja-g	0.45	0.017	185	0.06	0.54	0.014	255	0.091	0.55	0.027	264	0.114
Guarrati	0.68	0.085	162	0.19	0.71	0.075	279	0.095	0.72	0.077	292	0.080
LaAtalaya	0.56	0.021	187	0.08	0.69	0.037	295	0.053	0.70	0.040	309	0.060
LaCruzdeElias	0.60	0.024	202	0.13	0.63	0.022	296	0.067	0.70	0.004	309	0.051
LasArenas	0.56	0.029	209	0.14	0.77	0.031	308	0.071	0.71	0.033	319	0.058
LasBodegas	0.61	0.007	193	0.09	0.52	0.026	316	0.045	0.70	0.022	330	0.038
LasBrozas	0.72	0.007	206	0.05	0.58	0.010	314	0.050	0.74	0.012	327	0.066
LasEritas	0.62	0.019	196	0.20	0.65	0.037	305	0.121	0.63	0.020	318	0.103
LasTresRayas	0.74	0.035	193	0.09	0.60	0.048	276	0.059	0.73	0.045	284	0.066
LasVacas	0.59	0.023	194	0.06	0.57	0.016	286	0.063	0.70	0.021	300	0.092
LasVictorias	0.23	0.006	193	0.06	0.27	0.004	263	0.074	0.08	0.000	277	0.103
LlanosdeBoveda	0.57	0.026	201	0.11	0.63	0.024	292	0.059	0.66	0.027	306	0.047
Paredinas	0.67	0.009	188	0.07	0.41	0.004	281	0.013	0.81	0.012	289	0.144
Zamarron	0.66	0.011	197	0.05	0.61	0.019	288	0.083	0.79	0.014	301	0.100
平均值	0.55	0.022	194	0.098	0.58	0.023	287	0.074	0.63	0.023	299	0.087

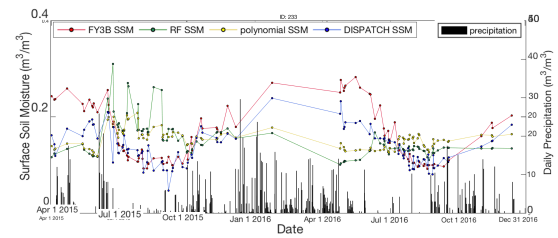
于每个气候站点,干早期由于受到蒸腾作用的影响保持较低SM值的水平。416站点干早期SM降尺度数据比FY3B SM数据有更合理的数值,SM的增加量也更大。总体来看,RF相比多项式拟合的方法降尺度数据能够更准确地反应时间变化,而DISPATCH降尺度结果与原始数据随时间变化趋势较为相似,主要归因于DISPATCH降尺度结果十分依赖于原始SM,不能较好地从时间和空间尺度表征SM变化细节。除了干早期之外,从2015年10月到2016年5月和2016年10月到2016年12月,由于较冷的气候和频繁降水,所有站点的土壤表面湿度都保持着较高的数值。

4 结论

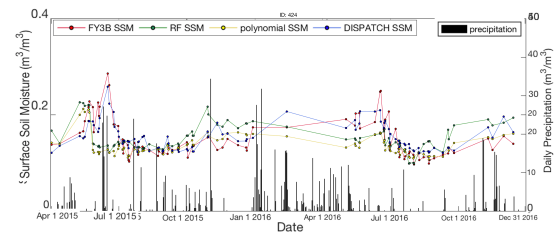
本文分别基于RF方法、多项式拟合法和DISPATCH方法,融合可见光、热红外和地表高程参量对FY3B微波土壤湿度数据进行降尺度,使其空间分辨率从25 km提高至1 km。根据FY-3B与相关输入数据源过境时间不匹配现象,将FY3B升降轨SM数据、Terra和Aqua卫星的MODIS产品重组为四组数据AM+Terra、AM+Aqua、PM+Terra和PM+Aqua,将四组数据分别应用于RF、多项式拟合和DISPATCH降尺度模型的建立,通过升降轨四组对照实验对比分析FY3B降尺度的最优数据组合。对比结果显示RF降尺度模型拟合效果最好,精度最高。此外,三种方法基于PM+Terra和PM+Aqua建立的降尺度模型模拟效果优于AM+Terra、AM+Aqua。

基于RF方法的四组数据降尺度结果显示Terra和Aqua降尺度结果差异较小。由于MODIS云污染和风云三号卫星扫描间隙会导致单一数据组降尺度结果存在大量无效值,本文基于四组数据建立合成数据以改善SM降尺度结果。降尺度结果显示多项式拟合降尺度结果低估了SM值,基于DISPATCH方法的降尺度合成数据具有明显的板块效应,像素存在较为明显的锯齿现象。直观地,基于RF方法的合成降尺度结果呈现较好的空间表达效果。

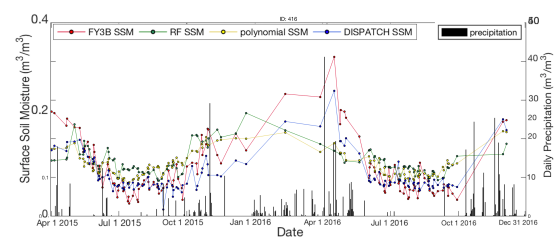
基于REMEDHUS土壤湿度原位站点数据和ECA&D气象站点数据对降尺度结果进行验证。验证过程主要包含两个部分:首先,基于RF方法对AM+Terra、AM+Aqua、PM+Terra和PM+Aqua四组数据的降尺度结果进行验证。验证结果显示基于PM+Terra数据的RF降尺度效果表现最好。之后,基于整个REMEDHUS土壤湿度原位站点数据和



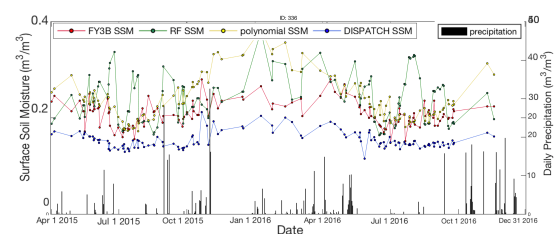
(a)



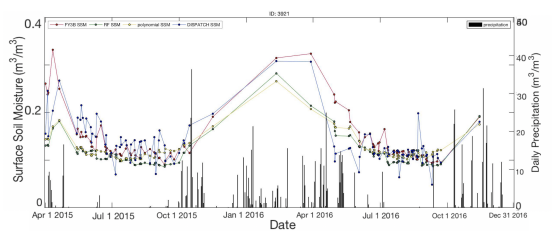
(b)



(c)



(d)



(e)

图18 降尺度前后结果与降水量关系图,(a)233站点,(b)424站点,(c)416站点,(d)336站点,(e)3921站点

Fig.18 The Diagram between downscaling results and original data, (a)No.233 station(b) No.424 station, (c)No.416 station, (d)No.336 station, (d)No.3921 station

ECA&D气象站点数据验证三种方法采用单一数据组的降尺度结果时空精度。基于原位实测数据验证结果显示RF和多项式拟合两种方法的降尺度结果与实测数据相关性大小相近,且都比DISPATCH方法的相关性大。三种方法的bVariance值大小相近,三种方法都能较好地表达空间SM数据的变异性。然而,基于RF方法的平均RMSD值最小,综合全部统计指标可以总结出基于RF的降尺度方法比多项式拟合与DISPATCH模型更为可靠,降尺度精度最高。从时间尺度上来看,DISPATCH方法降尺度结果与原始SM数据随时间变化几乎一致,但归因于降尺度结果依赖于原始数据,无法更精细地表征土壤湿度变化。RF与多项式拟合降尺度结果随时间变化趋势相似,但RF降尺度结果与降水时间序列一致性更好。三种方法的降尺度结果与地面气象站点都具有较好的时序一致性,干旱期和潮湿期能够在降尺度数据中较好地体现。

References

- [1] Givi J, Prasher S O, Patel R M. Evaluation of pedotransfer functions in predicting the soil water contents at field capacity and wilting point [J]. *Proceedings of The Fourth International Iran & Russia Conference*, 2004, **70**(2): 83–96.
- [2] Koster R D, Dirmeyer P A, Guo Z, et al. Regions of Strong Coupling Between Soil Moisture and Precipitation [J]. *Science*, 2004, **305**(0036–8075): 1138–1140.
- [3] Ochsner T E, Cosh M H, Cuenca R H, et al. State of the Art in Large-Scale Soil Moisture Monitoring [J]. *Soil Science Society of America Journal*, 2013, **77**(6): 1888–1919.
- [4] Robock A, Vinnikov K Y, Srinivasan G, et al. The Global Soil Moisture Data Bank [J]. *Bulletin of the American Meteorological Society*, 2000, **81**(6): 1281–1299.
- [5] Chauhan N S, Miller S, Ardanuy P. Spaceborne soil moisture estimation at high resolution: a microwave-optical/IR synergistic approach [J]. *International Journal of Remote Sensing*, 2003, **24**(22): 4599–4622.
- [6] LIN Li-Bin, BAO Yan-Song, ZUO Quan, et al. *Soil Moisture Retrieval over Vegetated Areas based on Sentinel-1 and FY-3C data*. REMOTE SENSING TECHNOLOGY AND APPLICATION, 林利斌, 鲍艳松, 左泉, 等. 基于Sentinel_1与FY_3C数据反演植被覆盖地表土壤水分. 遥感技术与应用) 2018, **33**(4): 750–758.
- [7] FAWWAZT. ULABY, GERALDA. BRADLEY, MYRONC. DOBSON. Microwave Backscatter Dependence on Surface Roughness, Soil Moisture, and Soil Texture: Part II—Vegetation-Covered Soil [J]. *IEEE TRANSACTIONS ON GEOSCIENCE ELECTRONICS*, 1979, **GE-17**(2): 286–295.
- [8] Njoku E G, Wilson W J, Yueh S H, et al. Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99 [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2002, **40**(12): 2659–2673.
- [9] Xiwu Z, Houser P R, Walker J P, et al. A method for retrieving high-resolution surface soil moisture from hydros L-band radiometer and Radar observations [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, **44**(6): 1534–1544.
- [10] Minet J, Verhoest N E C, Lambot S, et al. Temporal stability of soil moisture patterns measured by proximal ground-penetrating radar [J]. *Hydrology and Earth System Sciences Discussions*, 2013, **10**(4): 4063–4097.
- [11] Narayan U, Lakshmi V, Jackson T J. High-resolution change estimation of soil moisture using L-band radiometer and Radar observations made during the SMEX02 experiments [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, **44**(6): 1545–1554.
- [12] Piles M, Entekhabi D, Camps A. A Change Detection Algorithm for Retrieving High-Resolution Soil Moisture From SMAP Radar and Radiometer Observations [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2009, **47**(12): 4125–4131.
- [13] Wu X, Walker J P, Das N N, et al. Evaluation of the SMAP brightness temperature downscaling algorithm using active – passive microwave observations [J]. *Remote Sensing of Environment*, 2014, **155**: 210–221.
- [14] Hutengs C, Vohland M. Downscaling land surface temperatures at regional scales with random forest regression [J]. *Remote Sensing of Environment*, 2016, **178**: 127–141.
- [15] Montzka C, Jagdhuber T, Horn R, et al. Investigation of SMAP Fusion Algorithms With Airborne Active and Passive L-Band Microwave Remote Sensing [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, **54**(7): 3878–3889.
- [16] Wagner W, Dorigo W, Jeu R d, et al. FUSION OF ACTIVE AND PASSIVE MICROWAVE OBSERVATIONS TO CREATE AN ESSENTIAL CLIMATE VARIABLE DATA RECORD ON SOIL MOISTURE [J]. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2012, **1-7**: 315–321.
- [17] Chen N, He Y, Zhang X. NIR-Red Spectra-Based Disaggregation of SMAP Soil Moisture to 250 m Resolution Based on OzNet in Southeastern Australia [J]. *Remote Sensing*, 2017, **9**(1): 51.
- [18] 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: Earth and Environment*, 2009, **33**(2): 224–250.
- [19] Merlin O, Walker J, Chehbouni A, et al. Towards deterministic downscaling of SMOS soil moisture using MODIS derived soil evaporative efficiency [J]. *Remote Sensing of Environment*, 2008, **112**(10): 3935–3946.
- [20] Merlin O, Rudiger C, Al Bitar A, et al. Disaggregation of SMOS Soil Moisture in Southeastern Australia [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, **50**(5): 1556–1571.
- [21] K.YA.VINNIKOV, I.B.YERKEPOVA. Soil Moisture: Empirical Data and Model Results [J]. *Journal of Climate*, 1990, **4**: 14.
- [22] Srivastava P K, Han D, Ramirez M R, et al. Machine Learning Techniques for Downscaling SMOS Satellite Soil

- Moisture Using MODIS Land Surface Temperature for Hydrological Application [J]. *Water Resources Management*, 2013, **27**(8): 3127–3144.
- [23] suk L C, Eunha S, Dong P J, *et al.* Estimation of soil moisture using deep learning based on satellite data: a case study of South Korea [J]. *GIScience & Remote Sensing*, 2018, **56**(1): 43–67.
- [24] Kim S, Jeong J, Zohaib M, *et al.* Spatial disaggregation of ASCAT soil moisture under all sky condition using support vector machine [J]. *Stochastic Environmental Research and Risk Assessment*, 2018, **32**(12): 3455–3473.
- [25] Liu Y, Yang Y, Jing W, *et al.* Comparison of Different Machine Learning Approaches for Monthly Satellite-Based Soil Moisture Downscaling over Northeast China [J]. *Remote Sensing*, 2017, **10**(2): 31.
- [26] Zhao W, Sánchez N, Lu H, *et al.* A spatial downscaling approach for the SMAP passive surface soil moisture product using random forest regression [J]. *Journal of Hydrology*, 2018, **563**: 1009–1024.
- [27] Breiman L. Random Forests [J]. *Machine Learning*, 2001, **45**: 5–32.
- [28] Carlson T. An Overview of the "Triangle Method" for Estimating Surface Evapotranspiration and Soil Moisture from Satellite Imagery [J]. *sensors*, 2007, **7**: 1578–1611.