引文格式:李琼,张东,钟波,等.GRACE反演长江流域陆地水储量变化的机器学习降尺度方法研究[J].武汉大学学报(信息科学版),2026,51(6):1138-1155.DOI:10.13203/j.whugis20240310

Citation:LI Qiong,ZHANG Dong,ZHONG Bo,et al.Machine Learning Downscaling Methods for Inversion of Terrestrial Water Storage Anomalies in the Yangtze River Basin by GRACE[J].Geomatics and Information Science of Wuhan University,2026,51(6):1138-1155.DOI:10.13203/j.whugis20240310

GRACE反演长江流域陆地水储量变化的机器学习降尺度方法研究

李 琼[1,4] 张 东[1] 钟 波[2,3,4] 苏 勇[1] 杨翼宇[1]

1 西南石油大学土木工程与测绘学院,四川 成都,610500

2 武汉大学测绘学院,湖北 武汉,430079

3 地球空间环境与大地测量教育部重点实验室,湖北 武汉,430079

4 华中科技大学国家精密重力测量科学中心,湖北 武汉,430074

摘 要:重力恢复与气候实验(gravity recovery and climate experiment,GRACE)卫星时变重力场反演陆地水储量变化(terrestrial water storage anomaly,TWSA)的空间分辨率有限,制约了其在中小尺度区域水循环和气候变化研究中的应用潜力。当前,基于机器学习的降尺度方法已在提升GRACE TWSA空间分辨率方面取得了显著进展。然而,预测因子的合理选取及其对模型性能的影响机制,以及降尺度结果的精度评估与不确定性分析等问题,仍有待进一步深入研究。以长江流域TWSA的降尺度处理为例,对比传统的水文模型降尺度方法和人工神经网络(artificial neural network,ANN)、随机森林(random forest,RF)、极端梯度提升(extreme gradient boosting,XGBoost)3种机器学习模型,将GRACE反演的长江流域空间分辨率为1°×1°的TWSA数据降尺度到0.25°×0.25°和0.1°×0.1°。为了对不同降尺度方法进行评价,首先使用全球陆面数据同化系统水文模型的TWSA数据进行闭合模拟实验,评估不同降尺度方法的性能;然后对GRACE TWSA数据进行降尺度处理,并采用水文站水位数据评估不同降尺度方法的性能和结果。研究结果表明,机器学习降尺度模型的性能受到预测因子数量的影响,随着预测因子数量的增加,降尺度性能也会持续增强,3种机器学习模型的降尺度性能优劣也会发生改变。采用偏最小二乘回归分析模型得到重要性较高的前6个预测因子,即归一化植被指数、土壤湿度、降水量、气温、径流和U形风,在机器学习降尺度模型中取得较好的效果,增加更多的预测因子对降尺度模型性能的提升相对较小。RF和XGBoost降尺度模型的性能表现较好且非常接近,ANN和水文模型降尺度模型的效果略差。此外,水文模型的降尺度结果依赖水文模型与TWSA数据的相关性,机器学习降尺度模型能够更好地融合水文、气象和植被等多种数据的变化特征,从而恢复中小尺度区域精细的TWSA信号。

关键词:GRACE;时变重力场;长江流域;陆地水储量变化;机器学习;降尺度分析

中图分类号:P237 **文献标识码**:A **收稿日期**:2025-03-07

DOI:10.13203/j.whugis20240310 **文章编号**:1671-8860(2026)06-1138-18

Machine Learning Downscaling Methods for Inversion of Terrestrial Water Storage Anomalies in the Yangtze River Basin by GRACE

LI Qiong[1,4] *ZHANG Dong*[1] *ZHONG Bo*[2,3,4] *SU Yong*[1] *YANG Yiyu*[1]

1 School of Civil Engineering and Geomatics, Southwest Petroleum University, Chengdu 610500, China

2 School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China

3 Key Laboratory of Geospace Environment and Geodesy, Ministry of Education, Wuhan 430079, China

4 National Gravitation Laboratory, Huazhong University of Science and Technology, Wuhan 430074, China

Abstract: **Objectives**: The spatial resolution of the time-varying gravity field model provided by gravity recovery and climate experiment(GRACE) satellite gravity for retrieving terrestrial water storage anomaly

基金项目:国家自然科学基金(42374002,41931074,42374004);国家精密重力测量科学中心开放课题(NGL-2025-015,NGL-2025-016)。

第一作者:李琼,副教授,主要研究方向为卫星重力学与应用。qiongli@swpu.edu.cn

通信作者:钟波,博士,教授。bzhong@sgg.whu.edu.cn

(TWSA) is limited, which restricts its application potential in the study of regional water cycle and climate change. The current machine learning downscaling methods have been effective in improving the spatial resolution of GRACE TWSA data, but further exploration is needed on the reasonable selection of predictive factors and their impacts on the performance of machine learning models, as well as the accurate evaluation of downscaling results. **Methods:** A hydrological model downscaling method and three machine learning model downscaling methods, including random forest (RF), extreme gradient boosting (XGBoost), and artificial neural network (ANN), were adopted to downscale the TWSA data obtained from GRACE inversion of the Yangtze River Basin with the spatial resolutions from 1°×1° to 0.25°×0.25° and 0.1°×0.1°, respectively. To evaluate different downscaling methods, a closed-loop simulation experiment was conducted using TWSA data from global land data assimilation system hydrological model to evaluate the performance of different downscaling methods. Subsequently, the TWSA data obtained from GRACE was downscaled, and the performance and results of different downscaling methods were comprehensively evaluated using measured water level data. **Results:** The performance of machine-learning-models-based downscaling method is sensitive to the number of predictive factors. As the number of predictors increases, the overall downscaling performance improves, although the magnitude of improvement varies among models. Partial least squares regression analysis indicates that the six most important variables ,normalized difference vegetation index (NDVI), soil moisture, precipitation, temperature, runoff, and U-wind, are sufficient to achieve robust performance, while the inclusion of additional predictors yields only marginal gains. Closed-loop simulation experiments further reveal differences among models. Compared with RF and XGBoost models, ANN model shows relatively lower performance metrics; however, it produces the most favorable downscaled spatial patterns. When evaluated against water level observations in the Yangtze River Basin, TWSA remains consistent with observed water levels both before and after downscaling. In particular, RF-based results show a marked improvement in correlation, with all coefficients exceeding 0.7. Using Poyang Lake as a representative case, the long-term trends of the machine learning downscaling results were compared with those of the three predictors with the highest variable importance in projection scores, including NDVI, soil moisture, and precipitation. The results indicate that the downscaled outputs exhibit long-term trends highly consistent with NDVI and soil moisture, and their spatial variations closely correspond to the extent of water-covered areas in Poyang Lake. In contrast, precipitation does not display a similar trend pattern. **Conclusions:** The best downscaling methods are RF and XGBoost, while the downscaling methods of ANN and hydrological model perform poorly. Additionally, the results of hydrological model downscaling depend on the correlation between the hydrological model and GRACE data. However, machine learning downscaling methods can better integrate the changing characteristics of different auxiliary data such as hydrology, meteorology, and vegetation (especially important predictive factors), thus enabling better recovery of detailed TWSA signals in the watershed.

Key words: GRACE; time-varying gravity field; the Yangtze River Basin; terrestrial water storage anomaly (TWSA); machine learning; downscaling analysis

陆地水包括地下水、土壤湿度、地表水(河流、湖泊和水库)、雪水当量和植被冠层含水等各种形式的水体,是地球系统赖以维持的重要资源,也是水文气候变化的重要指标之一。陆地水储量变化(terrestrial water storage anomaly, TWSA)的准确估计不仅为水资源的科学监测和合理利用提供有效观测数据,还能为水文气候变化的研究与预测、洪旱灾害的预报预警等提供重要信息[1-2]。由于传统地面水文观测站存在单站覆盖范围小、站点分布不均匀等特点,很难获取空间上全覆盖的长期连续观测数据;不同的水文水资源模型虽然能够提供全球覆盖且长期连续的TWSA,但因观测数据的不均匀性以及采用模型、算法和因素侧重点不同,各类水文模型间提供的信息存在一定的差异[3]。重力恢复与气候实验(gravity recovery and climate experiment, GRACE)卫星及其后继卫星(GRACE Follow-On)的实施为中大空间尺度TWSA的长期连续

监测提供了有效技术手段,已在全球环境变化各领域的研究中发挥了重要作用,包括中大型流域地表水及地下水储量变化[4-7]、极端气候下的洪旱灾害事件[8-11]、极地和高山冰川消融[12-15]以气候变化和人类活动的响应研究等[16-19]。然而,利用GRACE时变重力场模型直接反演TWSA的空间分辨率较低,限制了其在中小尺度上的TWSA监测与应用。

近年来,为满足中小尺度区域TWSA的研究需求,国内外学者提出了多种降尺度方法来提高GRACE反演结果的空间分辨率。文献[20]最早提出了降尺度方法,包括基于模型的动力降尺度方法和基于数据的统计降尺度方法,该方法广泛应用于气候及气象模型的研究。在GRACE时变重力场反演TWSA的降尺度方法中,动力降尺度方法是将目标数据集整合至物理模型中产生高分辨率数据的数据同化方法,即通过水文模型将GRACE观测信息整合到陆地水储量的估算中[21]。这种方法对水文模型初始边界条件的依赖性强,且需要强大的计算能力[22]。统计降尺度方法建立较低分辨率被预测变量与高分辨率预测因子之间的统计关系,相比于动力降尺度方法,具有计算成本低、模型相对简单等特点[23]。

目前常用的统计降尺度方法是基于全球水文模型的降尺度方法,即采用线性回归方法,利用高时空分辨率的水文模型。例如利用全球陆地数据同化系统(global land data assimilation system,GLDAS)数据对GRACE反演的TWSA进行降尺度,得到对应的高时空分辨率GRACE TWSA数据,进而通过水量平衡方程将其应用于蒸散发量、地下水储量变化等水文学研究[24-26]。基于全球水文模型的统计降尺度方法是通过建立水文模型获取的TWSA与GRACE获取的TWSA之间的线性关系实现的,简单的线性模型不能准确且全面地描述GRACE TWSA与各水文要素之间的复杂关系[23]。机器学习模型能够更好地处理复杂的非线性问题,也可以捕捉GRACE反演的陆地水储量信息与水文变量之间的复杂关系。基于机器学习模型的统计降尺度方法已成为GRACE数据降尺度及相关应用研究的重要手段。文献[27]基于人工神经网络(artificial neural network,ANN)的非参数模型,将GRACE TWSA和GLDAS的水文变量进行集成,实现了GRACE TWSA的降尺度,并将其应用于地下水位的估算。文献[28]利用ANN模型以低分辨率GRACE和高分辨率的降水量、温度、坡度、土壤类型等变量作为输入数据,直接生成了美国加利福尼亚州地区约16 km^2的高分辨率地下水储量变化结果,并采用地面测井地下水位进行验证,结果表明,ANN模型降尺度方法可以很好地提高地下水储量的空间分辨率。文献[29]和文献[30]均采用随机森林(random forest,RF)方法将TWSA和地下水储量数据的空间分辨率从1°×1°降至0.25°×0.25°。文献[29]选取的预测因子为降水量、蒸散、径流、土壤湿度、雪水当量和植被冠层含水量6个变量,文献[30]除采用这6个水文变量外,还采用了数字高程模型、坡向、坡度、温度和风速等地理参数。引入欧洲中期天气预报中心(European centre for medium-range weather forecasts,ECMWF)为了第五代再分析(ECMWF Reanalysis v5,ERA5)土地部分(ERA5-Land)、植被指数(normalized difference vegetation index,NDVI)以及全球降水观测计划(global precipitation measurement,GPM)等数据作为变量,进一步提升降尺度的物理合理性与空间细节表现。文献[31]利用RF方法和极端梯度提升(extreme gradient boosting,XGBoost)模型将GRACE TWSA和地下水储量数据从空间分辨率1°×1°降至0.25°×0.25°,结果表明RF和XGBoost降尺度的结果相当。文献[32]采用XGBoost和ANN模型等机器学习模型将GRACE TWSA从空间分辨率1°×1°降至0.25°×0.25°,结果表明XGBoost模型的TWSA降尺度结果均优于ANN模型结果。现有研究表明,以ANN、RF、XGBoost等为代表的机器学习模型在现有GRACE TWSA的降尺度研究中,特别是在地下水储量、蒸散发量以及干旱指数等与GRACE TWSA具有强相关水文参数估计时,各类机器学习降尺度方法表现出较好的应用效果。但需要注意的是,机器学习方法在应用过程中可能会面临观测数据的可用性与不确定性、模型方法可解释性与可扩展性等问题[33-34]。机器学习模型的选取、预测因子的选取以及模型参数的调整等均会对机器学习模型的效果产生影响。因此,在GRACE TWSA降尺度及相关应用研究中,如何定量评估被预测变量与预测因子的相关性并合理选取预测因子,以及如何选取合适的机器学习模型并进行参数调整等问题,是目前机器学习降尺度方法应用研究中亟待解决的关键问题。

本文采用了ANN、RF和XGBoost 3种机器

学习模型降尺度方法,针对GRACE反演的长江流域TWAS数据进行降尺度处理,利用GLDAS水文模型及实测水位等数据对机器学习模型的性能及降尺度结果进行系统分析与评估。研究主要包括3个部分,首先利用多变量统计分析方法定量评估长江流域各类水文、气象和植被等预测因子对估算该区域TWSA的贡献量;然后利用不同空间分辨率的GLDAS水文模型设计数值模拟实验进行闭合验证,重点针对不同预测因子对3种机器学习模型性能及降尺度结果的影响进行分析与评估;最后利用3种机器学习降尺度方法对GRACE反演的长江流域TWSA进行降尺度计算,将降尺度前后的结果与长江流域实测水位数据进行统计分析,评估各类机器学习降尺度方法的可靠性。研究结果为GRACE时变重力场在流域或小尺度水文应用研究中机器学习降尺度方法的选取提供科学依据。

1　研究区域与数据

1.1　研究区概况

长江流域是中国第一大流域,地处90°33′E至122°25′E,24°30′N至35°45′N(见图1)。长江发源于青藏高原,自西南向东流,最终汇入东海。长江流域总面积约为180万 km²,干流全长约6 300 km,约占全国河流径流总量的36%。流域跨中国东部、中部和西部三大经济区。长江流域内地势复杂,呈多级阶梯地形,拥有丰富多样的气候特征。年内降水分布表现出不均匀性,冬季降水最少,春季逐渐增加,夏季降水高峰出现在6月—7月,秋季各地降水逐渐减少。多年平均陆面蒸发呈现出中下游大于上游、平原盆地大于山区、南岸大于北岸的分布趋势。年平均气温在不同地区呈现出东高西低、南高北低的特征。长江流域支流主要包括上游的金沙江、岷江、嘉陵江和乌江流域,中游的汉江、洞庭湖和鄱阳湖流域以及下游的太湖流域。长江流域的地理和气候特征直接塑造了其自然环境和生态条件,深入研究该地区的TWSA对生态环境保护至关重要。

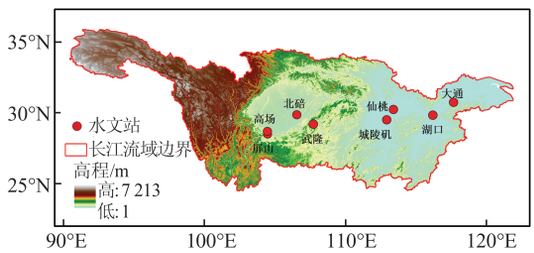


图1　长江流域位置及水文站示意图

Fig. 1　Location of the Yangtze River Basin and Schematic Diagram of Hydrological Stations

1.2　GRACE数据

本文使用德克萨斯大学空间研究中心(center for space research,CSR)发布的CSR Release-06 GRACE Level-2时变重力场模型(https://www2.csr.utexas.edu/grace/RL06.html),覆盖2002-04—2017-06共163个月的数据,其中20个月的数据缺失使用三次样条插值填补。将球谐系数模型转换为等效水高变化的过程中,进行的预处理步骤包括替换C_{20}项与一阶项、冰川均衡调整[35]、300 km高斯滤波和去相关滤波以及削弱高频噪声和南北条带误差[36-37]。在扣除多年平均值后,得到全球等效水高异常TWSA。最后利用正演模型对空间滤波后的数据进行泄漏误差改正[37]。上述处理减弱了系统误差及滤波导致的信号衰减与泄漏。在此基础上,利用时变重力场模型反演全球地表质量变化,等效水高变化ΔH可表示为[38]:

$$\Delta H(\theta,\lambda)=\frac{a\rho_{ave}}{3\rho_w}\sum_{l=0}^{\infty}\sum_{m=0}^{l}\bar{P}_{l,m}\cos\theta\frac{2l+1}{1+k_l}(\Delta\bar{C}_{l,m}\cos(m\lambda)+\Delta\bar{S}_{l,m}\sin(m\lambda)) \tag{1}$$

式中,a为地球赤道半径;ρ_{ave}为地球的平均密度;ρ_w为水的平均密度(1 000 kg/m³);θ和λ分别为地心余纬和地心经度;l和m分别为球谐系数的阶、次数;$\bar{P}_{l,m}$为完全规格化的缔合勒让德函数;k_l为l阶负荷勒夫系数;$\Delta\bar{C}_{l,m}$和$\Delta\bar{S}_{l,m}$为球谐系数的变化量,由各个月时变重力场模型的球谐系数扣除研究时间段内的平均值获得。

1.3　水文气象数据

1.3.1　GLDAS数据

GLDAS是由美国戈达德空间飞行中心和美国国家环境预报中心合作开发的一个陆面模型系统(https://disc.gsfc.nasa.gov/datasets keywords=GLDAS),旨在融合卫星和地面观测数据,生成全球陆面状态及通量产品,GLDAS包括5种陆面模型。本文选用GLDAS模型的两种TWSA,一种是Noah 2.1模型的土壤湿度、雪水当量和植物冠层含水量的总和作为TWSA,距平值记为Noah TWSA;另一种是汇流域陆面模型(catchment land surface model,CLSM)2.2的TWSA数据,距平值记为CLSM TWSA。需注意的是,CLSM 2.2模型数据起点时间为2003-02-01,与GRACE数据不完全同步。

1.3.2 ERA5-Land数据

ERA5-Land数据提供了从1950年至今的长时间序列数据,并持续进行更新,以满足土地监测应用的需求(https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview)。该数据将模型输出与全球原位观测数据结合,提供了0.1°×0.1°的高分辨率网格数据。本文选用了ERA5-Land提供的月平均再分析数据作为机器学习降尺度模型的预测因子。

1.3.3 GPM降水量与NDVI数据

GPM是一项国际卫星任务,旨在更精确地估算地球大部分表面的降水量,采用GPM多卫星集成反演(integrated multi-satellite retrievals for GPM,IMERG)算法,整合GPM卫星中的所有微波仪器的数据,生成GPM 3级产品。本文使用IMERG 06版本3级产品月时间间隔的数据(https://disc.gsfc.nasa.gov/datasets keywords=GPM),空间分辨率为0.1°×0.1°。

NDVI是一种用于反映大范围区域内植被生长和覆盖情况的重要指标。本文使用Terra卫星的月植被指数产品MOD13C2 6.1的NDVI数据(https://lpdaac.usgs.gov/products/mod13c2v006/),空间分辨率为0.05°×0.05°。

1.3.4 水文站数据

本文使用了长江流域8个水文站的地面实测水位日值,水文站的分布如图1所示。8个水文站分别是长江7个主要支流的控制站,观测数据时间为2002-01—2013-12。为了与GRACE模型及其他预测因子的时间分辨率一致,首先将这些日水位数据转化为月尺度数据,随后计算每个站点月平均水位的距平值,用以描述该月水位相对于长期平均水位的偏差。这8个水文站实测水位变化可视为各支流总的水位变化。因此,在使用实测水位进行评估时,实测水位变化将与GRACE计算的各流域TWSA数据进行比较和分析,水文站及其对应的流域和集水面积如表1所示。

表1 长江主要支流水文站信息

Table 1 Information of Hydrological Stations Along Major Tributaries of the Yangtze River

流域	水文站	集水面积/km²	流域	水文站	集水面积/km²
岷江	高场	135 378	汉江	仙桃	144 683
金沙江	屏山	458 592	洞庭湖	城陵矶	264 712
嘉陵江	北碚	156 736	鄱阳湖	湖口	162 225
乌江	武隆	83 035	长江	大通	1 705 383

为了确保水文测站实测水位数据和GRACE TWSA数据的一致性,常通过实测水位数据与比产量值相乘来实现。本文通过计算实测水位数据和GRACE获取的1°格网TWSA数据两者间最小的均方根误差(root mean square error,RMSE)确定最佳K值,用来替代比产量值,最佳K值和水位变化的TWSA计算公式如下:

$$K^{*}=\arg\min_{K}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(G_i-K\times T_{L,i})} \quad (2)$$

$$T_{\mathrm{TWSA},i}=T_{L,i}\times K^{*} \quad (3)$$

式中,K^{*}表示最佳K值;n表示数据点的数量;G_i表示第i个水文站GRACE卫星反演的TWSA数据;$T_{\mathrm{TWSA},i}$表示水位变化的TWSA;$T_{L,i}$表示第i个水文站观测到的水位变化数据;$\arg\min_{K}$表示在这个函数中,找到使该函数取得最小值的K值。

2 原理与方法

2.1 偏最小二乘回归模型预测因子选择

在GRACE降尺度研究中,预测因子的选取对降尺度模型的性能和计算结果都至关重要,预测因子的选取主要依赖于经验法。本文在参考前人研究的基础上,选择了12个水文、气象因子,包括降水量、NDVI、土壤湿度、土壤温度、径流、蒸发、地面气压、U形风、V形风、露点温度、空气温度和雪水当量[29-30,39-40](表2)。

表2 本文数据来源及其空间分辨率

Table 2 Data Sources Used in This Article and Their Spatial Resolutions

数据名称	数据来源	空间分辨率
GRACE TWSA	CSR Release-06 GRACE Level-2	1°×1°
土壤湿度、雪水当量、径流等	GLDAS Noah 2.1/GLDAS CLSM 2.2	1°×1°/0.25°×0.25°
降水量	GPM	0.1°×0.1°
NDVI	MOD13C2 6.1	0.05°×0.05°
土壤湿度、U形风、V形风、径流等	ERA5-Land	0.1°×0.1°

为了准确地评估TWSA的主导因素,本文采用了一种多变量统计分析方法,即局部最小二乘回归(partial least squares regression,PLSR)模型,计算12个水文、气象因子的变量重要性指标(variable importance in projection,VIP)得分,定量评估每个水文、气象因子对TWSA的贡献量。基于PLSR权重的加权平方和计算VIP得分[22]的公式为:

$$V_j=\left[\frac{p\sum_{m=1}^{h}R(Y,t_m)w_{m,j}^2}{\sum_{m=1}^{h}R(Y,t_m)}\right]^{\frac{1}{2}} \tag{4}$$

式中,h代表提取的成分数量;p代表预测变量的数量;t_m表示第m个分量;$R(Y,t_m)$表示第t_m个成分对于被预测变量Y的解释能力;$w_{m,j}$代表预测变量对于被预测变量的边际贡献。一般认为VIP得分大于0.8的预测变量对被预测变量具有显著性影响。本文利用PLSR计算预测因子的VIP得分,从而选择VIP得分较高的水文、气象因子作为预测因子输入到机器学习降尺度模型中。

2.2　机器学习降尺度方法

本文选取了ANN、RF和XGBoost 3种机器学习模型。ANN是一种常见的机器学习模型,模拟了人脑的神经处理能力,由输入层、隐藏层和输出层组成,常用于处理各种复杂非线性问题[41],其优点主要是无需对底层过程的函数形式进行先验假设,能够建立复杂非线性数据之间的关系,并且有助于降低数据中的噪声,但在处理复杂问题时容易出现过拟合[42]。RF则适用于高维数据和强非线性问题,能够捕捉变量之间的非线性关系[43],主要是通过自动随机选择训练样本来生成多个独立的回归树,每个回归树都是使用自举采样方法生成的样本,最终的预测结果是所有树的输出结果的平均值。RF生成回归树之间的独立性和低相关性,可以有效避免过拟合问题的发生[44]。XGBoost是一种结合了可扩展增强算法和集成机器学习技术的方法,该算法通过多次迭代将多个弱学习器组合成一个更强、更准确的学习器[45]。XGBoost被认为是传统的梯度提升决策树模型的改进模型,其在拟合决策树模型时关注的是残差而不是结果,可有效地克服过拟合问题[46]。从现阶段降尺度应用研究中可以看出,虽然这3种机器学习方法存在各自优缺点,但均在一定程度上满足提高空间分辨率的需求。本文以这3种机器学习模型为代表,将其应用于长江流域GRACE TWSA的降尺度并进行评估。

机器学习降尺度方法的一般过程是将GRACE获取的较低分辨率(如1°×1°)的TWSA与选取的预测因子在该空间分辨率下进行训练,基于该训练模型和高分辨率的预测因子计算高分辨率(0.25°×0.25°)下的TWSA。为了保证GRACE TWSA数据在降尺度前后的一致性,对模型预测的0.25°×0.25°的TWSA进行残差改正,使得高分辨率下的TWSA平均值与降尺度前的TWSA平均值一致。以1°×1° GRACE TWSA降尺度到0.25°×0.25°为例,机器学习降尺度方法的具体步骤如下:

1)预测因子选择和预处理。选取对GRACE TWSA降尺度可能有贡献的水文、气象及植被因子等作为预测因子并获取相应的高分辨率数据,剔除异常值并对缺失值进行填补,将预测因子的数据重采样至1°×1°和0.25°×0.25°。

2)机器学习模型训练。利用1°×1°下的预测因子和GRACE TWSA数据建立机器学习模型,将各类数据划分为训练集数据(75%)和测试集数据(25%),进行参数调整并验证模型性能,最终建立训练好的机器学习模型并获得1°×1°下机器学习模型预测结果。

3)残差改正量的计算。计算1°×1°下机器学习模型预测结果和实际的GRACE TWSA数据之间的差异,保留1°×1°残差改正量。利用克里金插值方法将1°×1°的残差改正量插值为0.25°×0.25°,获得0.25°×0.25°残差改正量,与最终降尺度后的数据分辨率相匹配。

4)降尺度预测与残差改正以0.25°×0.25°的预测因子作为输入参数,利用训练好的机器学习模型,构建0.25°×0.25°下机器学习模型预测结果,将获得的0.25°×0.25°残差改正量加回到该预测结果中,最终获取0.25°×0.25°的GRACE TWSA降尺度结果。

2.3　水文模型降尺度

本文采用了一种广泛应用的基于水文模型的统计降尺度方法,简称为水文模型(下文用SWMX表示)[24]。SWMX可在降尺度前后,确保每个1°×1°格网的陆地水储量异常值总和保持不变。SWMX具体计算公式为:

$$T_{d,i}=T_{\mathrm{GLDAS}_{0.25},i}+\frac{(T_{\mathrm{GRACE},i}-T_{\mathrm{GLDAS},i})\times A\times T_{\mathrm{GLDAS}_{0.25},i}}{\sum(T_{\mathrm{GLDAS}_{0.25},i}\times a_i)} \tag{5}$$

式中,$T_{d,i}$表示降尺度后的0.25°×0.25°格网数据;i表示月份;$T_{\mathrm{GRACE},i}$表示GRACE卫星反演得到1°×1°格网的TWSA数据;$T_{\mathrm{GLDAS},i}$和$T_{\mathrm{GLDAS}_{0.25},i}$分别表示SWMX GLDAS降尺度所用1°×1°和0.25°×0.25°的TWSA数据;A为1°×1°格网总面积;a_i为0.25°×0.25°格网的面积。

2.4　参数评估

为了选择最优的机器学习降尺度方法和对降尺度方法进行评估,使用了均方根误差(root

mean square error,RMSE)、相关系数(R)、纳什-萨特克利夫效率系数(Nash-Sutcliffe efficiency coefficient,NSE)3个评价指标,评价不同降尺度方法的性能。3个评价指标的计算公式为:

$$E_{\mathrm{RMSE}}=\sqrt{\frac{\sum_{i=1}^{N}(X_i-Y_i)^2}{N}} \tag{6}$$

$$R=\frac{\sum_{i=1}^{N}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i-\bar{X})^2}\sqrt{\sum_{i=1}^{N}(Y_i-\bar{Y})^2}} \tag{7}$$

$$E_{\mathrm{NSE}}=1-\frac{\sum_{i=1}^{N}(Y_i-X_i)^2}{\sum_{i=1}^{N}(X_i-\bar{X})^2} \tag{8}$$

式中,X_i和Y_i分别表示第i个真值和预测值的观测值;$\bar{X}$和$\bar{Y}$分别对应其平均值;N为样本总数。

综合上述基本原理与方法,本文降尺度数据处理过程与性能评估的流程如图2所示,包括3个部分:(1)数据准备与预处理。计算并重采样获取相应分辨率下的各类数据,包括GRACE TWSA、GLDAS、水文气象及植被数据。(2)降尺度结果获取。首先,利用PLSR模型定量分析预测因子的重要性并进行预测因子选取。然后采用3种机器学习降尺度模型获取GRACE TWSA降尺度结果,同时利用SWMX获取相同分辨率的降尺度结果。(3)统计分析不同模型获取的降尺度结果。将其与长江流域及二级流域实测水位数据进行验证和评估。

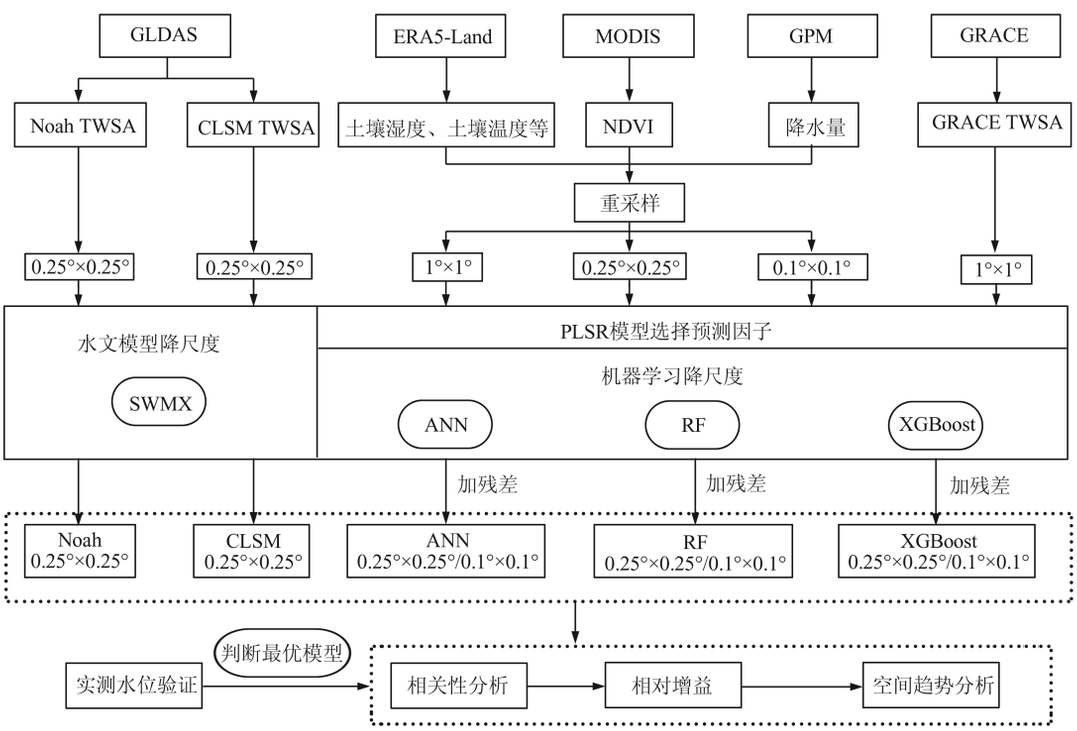


图2 降尺度数据处理流程图

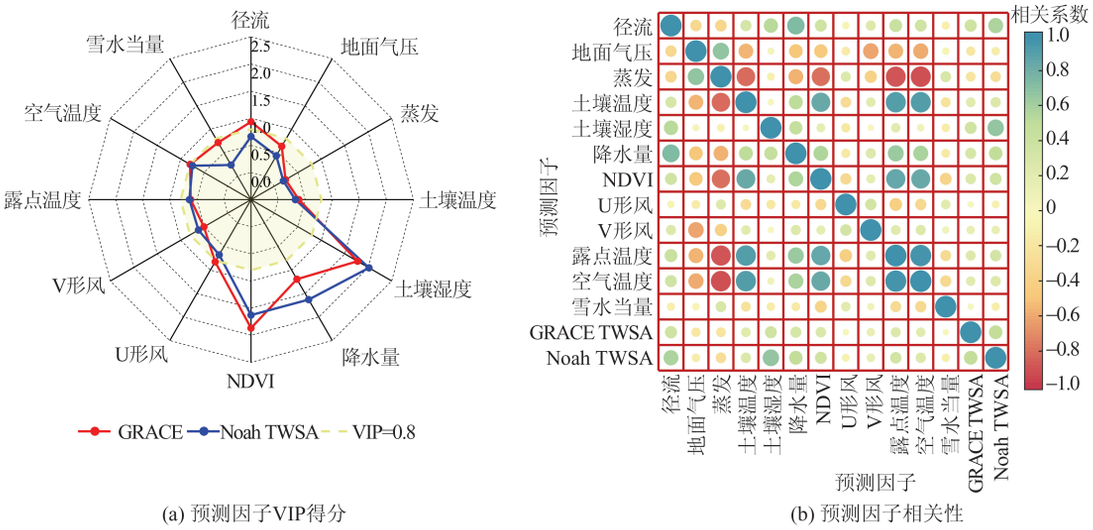
Fig. 2 Downscaling Data Processing Flowchat

3 结果与分析

3.1 预测因子选取

为了更深入地探究机器学习降尺度模型中12个预测因子对被预测因子的影响,计算的VIP得分如图3所示。NDVI、土壤湿度、降水量的VIP得分均大于1.0,表明这些因子在GRACE TWSA预测因子中具有较高的重要性。此外,径流、U形风、空气温度的VIP得分高于0.8,也是重要的预测因子。相比之下,其余预测因子的VIP得分相对较低,均小于0.8。同时,当GLDAS Noah获取的TWSA作为被预测因子时,本文也计算了12个预测因子的VIP得分,计算出的预测因子重要性排序与GRACE TWSA基本相同。NDVI、土壤湿度、降水量仍为VIP得分最高的预测因子,且VIP得分均大于1.0。但此时土壤湿度的VIP得分略高于NDVI,其原因是VIP得分计算时采用的土壤湿度为GLDAS模型数据,即直接是Noah TWSA的主要组成部分。

图3(b)给出了预测因子之间的相关性热图。



(a) 预测因子VIP得分　　(b) 预测因子相关性

图 3　不同预测因子 VIP 得分与相关性热图

Fig. 3　VIP Scores and Correlation Heatmap of Different Predictors

GRACE TWSA 与 Noah TWSA 具有很强的相关性。同时相关性最高的预测因子降水量、土壤湿度和 NDVI,也分别是 VIP 最高的预测因子。并且 GRACE TWSA 和 Noah TWSA 与其余预测因子的相关性基本一致,这也是两者在不同预测因子上的重要性排序相似的原因。

3.2　闭合模拟实验与结果分析

考虑到缺乏长期连续的高精度数据模型对降尺度后的 GRACE TWSA 进行系统的定量评估,GLDAS 则提供了多种时空分辨率的全球水文模型,因此采用 GLDAS 发布的 1°×1° Noah TWSA 作为被预测因子。选取相同的预测因子和降尺度方法获取 0.25°×0.25°的 TWSA,将其与 GLDAS 发布的 0.25°×0.25° Noah TWSA 进行对比分析,验证不同机器学习降尺度模型的有效性。图 4 给出了 2002-04—2017-06 时间范围内 GRACE 反演获得的长江流域 TWSA 以及 Noah 和 CLSM 两个陆面模型下 TWSA 随时间的变化情况。从图 4 可以看出,3 组 TWSA 的季节性变化和长期趋势特征基本一致。Noah 和 CLSM 的 TWSA 与 GRACE TWSA 的 R 分别为 0.86 和 0.71,有较强的一致性,表明可用 GLDAS 模型 TWSA 代替 GRACE TWSA 进行闭合模拟实验。

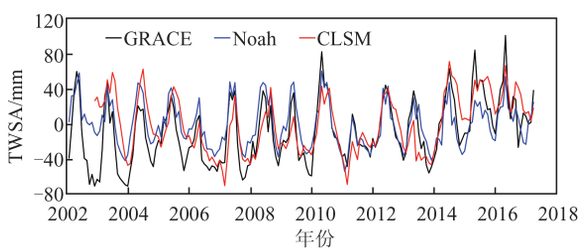


图 4　长江流域 TWSA 时间序列图

Fig. 4　Time Series of TWSA over the Yangtze River Basin

为了有效评估 3 种机器学习模型方法和相应的降尺度结果,从模型性能、降尺度性能和总残差 3 个方面进行了对比分析。模型性能指的是测试集上的性能;降尺度性能指的是降尺度结果与真值的一致性,均采用 R 和 RMSE 来衡量;总残差表示机器学习模型在 1°×1°格网下的预测值与原始值的残差绝对值均值。图 5 给出了随着预测因子数量的变化,模型性能、降尺度性能以及总残差的变化情况,其中 1～12 个预测因子是按照 PLSR 模型计算的 VIP 得分从高到低依次输入。图 5(a)显示,随着预测因子数量的增加,3 种机器学习模型的性能逐渐增强,但 3 种模型性能的相对大小会发生改变。当预测因子少于 4 个时,ANN 模型性能最好,RF 模型性能最差。随着预测因子数量继续增加,XGBoost 模型性能最好,ANN 模型性能最差。从图 5(b)中可知,随着预测因子数量的增加,只有 RF 模型的降尺度结果的 R 逐渐增强、RMSE 逐渐降低,即降尺度性能逐渐提高。当预测因子数大于 6 个时,降尺度性能提高放缓,XGBoost 模型的降尺度性能基本保持不变,ANN 模型的降尺度性能存在一定程度的下降。总体而言,ANN 模型的降尺度性能优于 RF 和 XGBoost 模型,这与模型性能的表现存在差异。图 5(c)显示了 3 种机器学习模型的总残差,总残差随着预测因子数量增加而逐渐减小。表明机器学习模型一定程度上是通过提高模

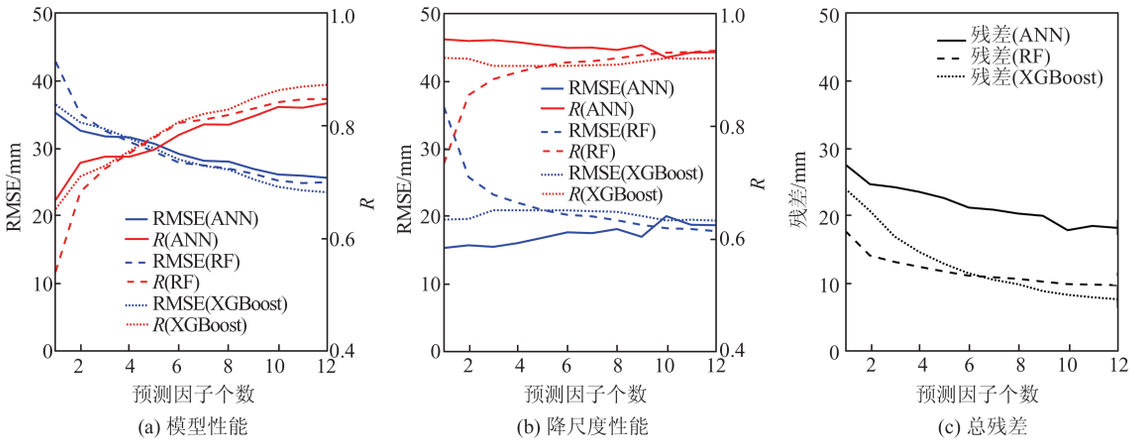


图5 不同预测因子数量对机器学习降尺度的影响

Fig. 5 The Impact of Different Numbers of Predictors on Machine Learning Downscaling

型性能和减小残差来提高降尺度性能,并且随着预测因子数量的增加,总残差的减小逐渐变缓。

为了进一步探讨预测因子数量(n)对预测结果的影响,本文针对RF模型,模拟预测因子数量为1~8个时2002-04的0.25°×0.25° TWSA。图6(a)~6(d)结果显示,预测因子数量为1~4个时,降尺度结果存在明显差异;图6(e)~6(h)结果显示,预测因子数量大于4个时,降尺度结果基本一致,且与图6(i)中0.25°×0.25°的Noah TWSA更接近。研究结果表明,利用PLSR模型评估VIP得分排名前6的预测因子进行机器学习模型降尺度时,可以获取稳定、可靠的结果。

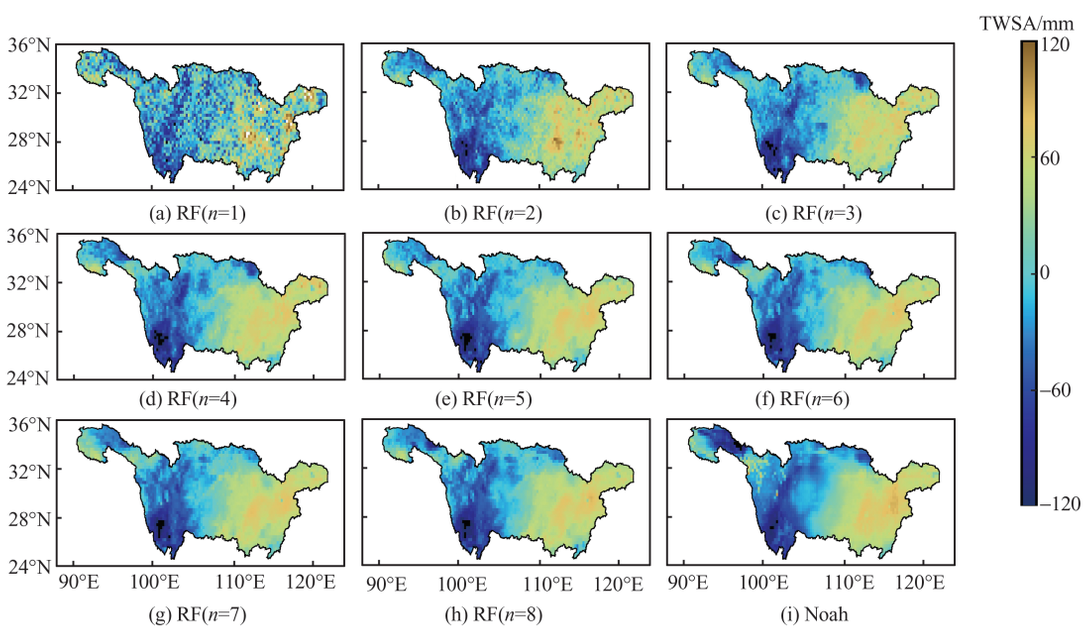


图6 不同预测因子数量下RF模型的TWSA降尺度结果

Fig. 6 TWSA Downscaling Results of RF Model with Different Numbers of Predictors

图7展示了3种机器学习方法均使用VIP得分最高的6个预测因子进行降尺度预测时,在测试集上的性能表现。从测试集的结果来看,3种模型性能相当,RF模型略优于XGBoost模型,ANN模型性能略低。此外,利用0.25°×0.25°的Noah TWSA数据作为真实值对降尺度结果进行评估,统计结果如表3所示。结果表明,在测试集上性能表现最差的ANN模型的降尺度结果优于其他各类方法,且SWMX降尺度方法的结果较差。考虑所有的预测因子的情况下,ANN在降尺度方面取得了最佳效果,其次是RF和XGBoost,而SWMX的降尺度效果最差。然而,从图5(a)、5(b)中可知,随着预测因子数量的增加,ANN的模型性能逐渐提升,降尺度性能逐渐下降。这意味着,ANN在仅使用1个预测因子,即模型性能最差时,降尺度效果最好。在图5(c)中可以看

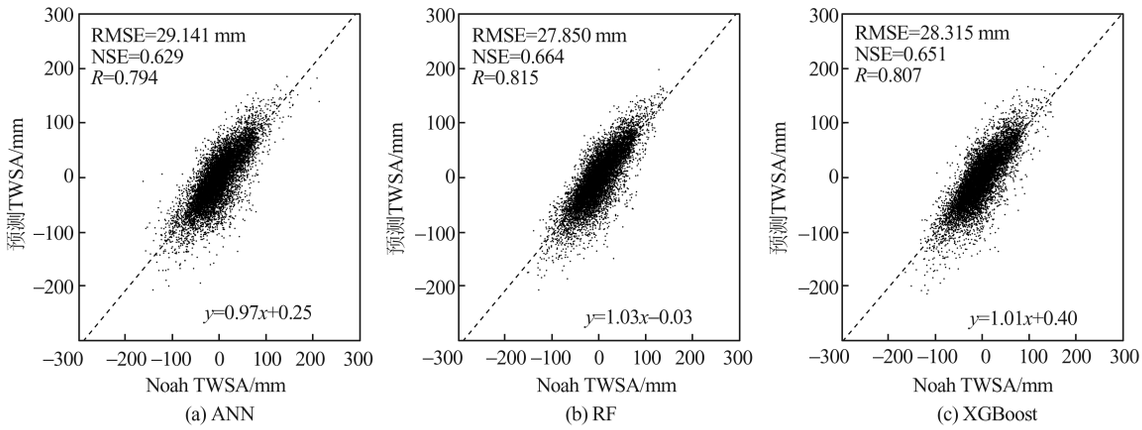


图 7　3种机器学习模型降尺度测试集性能

Fig. 7　Test Set Performance of Three Machine Learning Models for Downscaling

表 3　不同降尺度方法TWSA与 0.25°×0.25°Noah TWSA 的统计分析

Table 3　Statistical Analysis of Results from Different Models and 0. 25°×0. 25° Noah TWSA

模型	*R*	NSE	RMSE/mm
Noah(1°×1°)	0.966	0.932	12.869
SWMX	0.908	0.795	22.291
ANN	0.938	0.873	17.538
RF	0.913	0.833	20.130
XGBoost	0.906	0.822	20.791

出,ANN使用1个预测因子时将获得最大总残差。通过1°×1°的预测因子生成1°×1°的GRACE TWSA时模型残差最大,将该残差加回到生成的0.25°×0.25°的GRACE TWSA中,以保证降尺度前后相同网格中的总水储量保持不变,在总残差最大时尽可能多地保留原分辨率下的信息。这解释了在模型性能最差的情况下,ANN取得最佳的降尺度效果。ANN在模型性能最差的情况下引入了最多的残差,使其降尺度后更接近于原始1°×1°的Noah TWSA数据,也更接近0.25°×0.25°的Noah TWSA数据。该现象出现的原因可能是由于GLDAS Noah数据提供的TWSA与实际观测值之间存在差异。各个预测因子大多来源于实际观测数据,在机器学习降尺度过程中不可避免地引入了观测信息,使得降尺度结果与0.25°×0.25°的GLDAS TWSA数据之间产生差异。这种差异在一定程度上反映了0.25°×0.25°下Noah TWSA与真实观测之间本身存在的偏差。因此,基于GLDAS TWSA作为参考所评估得到的模型性能下降,未必能够真实反映降尺度结果的实际精度。为了更加客观地评估模型的降尺度能力,有必要进一步利用GRACE获取的TWSA数据,对机器学习降尺度结果开展独立验证与分析。

3.3　GRACE降尺度结果分析

利用CSR Release-06 GRACE Level-2时变重力场模型反演2002-04—2017-06时间范围内1°×1°的长江流域TWSA,以0.25°×0.25°和0.1°×0.1°的各类水文、气象和NDVI等要素作为预测因子,采用ANN、RF和XGBoots 3种机器学习模型将GRACE TWSA数据进行降尺度。同时,采用SWMX基于CLSM和Noah两种陆面模型计算的TWSA,将GRACE TWSA降尺度到0.25°×0.25°,通过统计分析评估降尺度结果并将其与流域实测水位数据进行验证。

3.3.1　模型性能和降尺度性能分析

图8给出了随着预测因子数量的变化,3种机器学习模型对长江流域GRACE TWSA进行降尺度处理时的模型性能、降尺度性能以及总残差的变化情况。采用的12个预测因子同样是按照PLSR方法计算的VIP得分从高到低依次输入,与数值模拟实验不同,此处降尺度性能采用8个水文站实测水位数据来进行评估。需要指出的是,为了便于与GRACE TWSA结果及降尺度结果对比统计分析,此处及下文采用的实测水位数据均为通过式(2)和式(3)转换后的水位变化的TWSA表示。

图8和图5对比显示,3种机器学习模型性能、降尺度性能及总残差随着预测因子的数量变化特征基本一致,其主要不同体现在3种机器学习模型的降尺度性能差异上,数值模拟分析结果表明,机器学习模型性能最差的ANN的降尺度性能最优。而采用实际观测数据的降尺度结果表明,机器学习模型性能和降尺度性能基本一致,即RF和XGBoost的模型性能相当,且明显

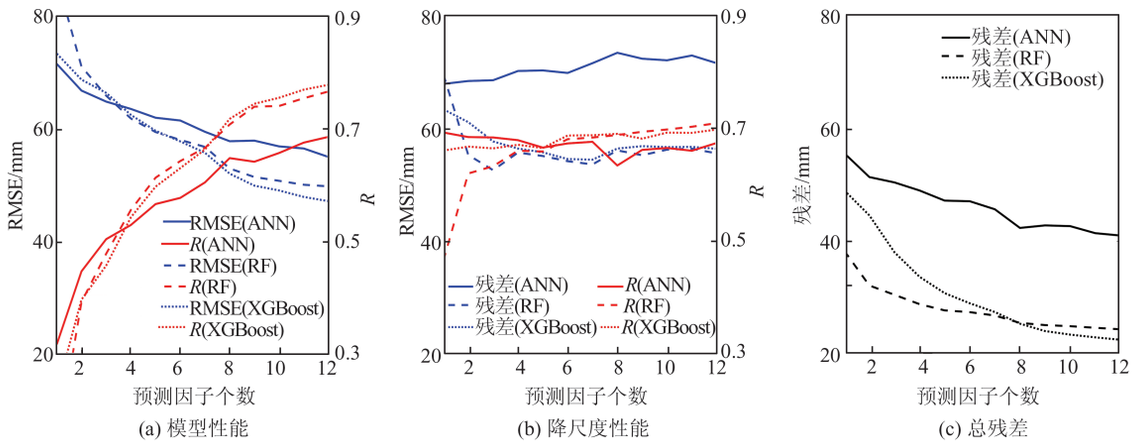


图8 长江流域不同预测因子数量对机器学习模型降尺度的影响

Fig. 8 Impact of the Number of Predictors on Machine Learning Model Downscaling in the Yangtze River Basin

优于ANN模型,这也能在一定程度上解释数值模拟实验中ANN方法取得最佳降尺度性能的可能原因是加回的总残差最大,使得降尺度结果更接近0.25°×0.25°的Noah TWSA。此外,当预测因子数量大于6个时,随着预测因子数量的增加,RF和XGBoost的模型性能和降尺度性能没有明显提升。因此,本文后续将使用前6个预测因子进行GRACE TWSA降尺度研究。

3.3.2 降尺度结果的相关性分析

表4和表5分别给出了降尺度前后在长江流域及其7个支流TWSA与实测水位的R和RMSE。GRACE TWSA在长江流域与实测水位相关性最高(R=0.824)。除嘉陵江流域外,其余6个支流的R均大于0.5,其中3个支流超过0.7,表明GRACE在大中尺度流域具有较好的适用性,而在中小尺度流域,其精度仍受空间分辨率限制。与降尺度前相比,SWMX和ANN模型降尺度后对结果提升有限,部分流域虽相关性变化不显著,但RMSE有所增加,如岷江、金沙江和鄱阳湖。相比之下,RF和XGBoost模型显著提高了R并降低了RMSE,两者性能接近,表现出较好的稳定性与一致性,ANN模型整体性能较弱。此外,从RMSE来看,0.1°×0.1°分辨率下的机器学习模型降尺度结果整体略优于0.25°×0.25°分辨率。

表4 长江流域降尺度前后TWSA与实测水位相关系数

Table 4 R Between GRACE TWSA and Measured Water Levels in the Yangtze River Basin

模型	分辨率	岷江	金沙江	嘉陵江	乌江	汉江	洞庭湖	鄱阳湖	长江
GRACE	1°×1°	0.650	0.715	0.115	0.683	0.501	0.734	0.745	0.824
SWMX CLSM	0.25°×0.25°	0.633	0.748	0.173	0.655	0.549	0.703	0.795	0.847
SWMX Noah	0.25°×0.25°	0.629	0.740	0.160	0.662	0.553	0.712	0.799	0.847
ANN	0.25°×0.25°	0.635	0.730	0.160	0.660	0.553	0.688	0.789	0.829
ANN	0.1°×0.1°	0.642	0.719	0.149	0.656	0.551	0.681	0.786	0.819
RF	0.25°×0.25°	0.826	0.885	0.398	0.726	0.672	0.758	0.789	0.896
RF	0.1°×0.1°	0.834	0.879	0.395	0.729	0.675	0.760	0.786	0.895
XGBoost	0.25°×0.25°	0.824	0.898	0.352	0.703	0.648	0.754	0.784	0.890
XGBoost	0.1°×0.1°	0.841	0.860	0.347	0.706	0.645	0.753	0.781	0.886

图9为长江流域及7个支流降尺度前、XGBoost和SWMX Noah模型降尺度到0.25°×0.25°以及水文站实测水位的时间序列变化。由于2012-06屏山水文站监测区域内的向家坝水电站开始蓄水,导致后续水位存在阶跃变化,该变化不能代表金沙江流域的TWSA,因此在下文的分析中,金沙江流域屏山水文站仅使用2012-06之前的数据。从图9可以看出,除了嘉陵江流域以外,长江流域及其他6个支流的降尺度前后的TWSA与实测水位变化均有较好的一致性,且采用XGBoost模型降尺度后与实测水位数据的一致性明显提高,而采用SWMX Noah降尺度后与原始数据仅存在微小差异。

3.3.3 降尺度结果的空间增益与时间增益

通过相对增益对降尺度前后的结果及降尺度方法的性能进行评估,该方法不仅可以衡量不

表 5　长江流域降尺度前后 TWSA 与实测水位 RMSE/mm

Table 5　RMSE Between GRACE TWSA and Measured Water Levels in the Yangtze River Basin/mm

模型	分辨率	岷江	金沙江	嘉陵江	乌江	汉江	洞庭湖	鄱阳湖	长江
GRACE	1°×1°	29.23	20.60	38.38	59.78	50.49	62.23	57.06	28.55
SWMX CLSM	0.25°×0.25°	29.86	24.14	33.60	57.44	48.54	62.67	57.50	25.32
SWMX Noah	0.25°×0.25°	30.03	24.07	32.56	56.86	47.41	62.33	56.93	25.07
ANN	0.25°×0.25°	30.06	24.89	33.57	58.27	49.28	64.00	58.84	26.57
ANN	0.1°×0.1°	30.17	25.00	33.78	57.92	49.78	63.66	58.89	26.65
RF	0.25°×0.25°	19.74	20.59	24.22	45.52	32.21	51.50	49.93	21.16
RF	0.1°×0.1°	19.40	20.09	24.04	44.21	31.97	50.44	49.28	20.95
XGBoost	0.25°×0.25°	19.86	20.04	25.14	44.59	32.72	51.68	49.10	20.93
XGBoost	0.1°×0.1°	19.56	20.94	24.97	44.31	32.66	50.35	48.23	20.51

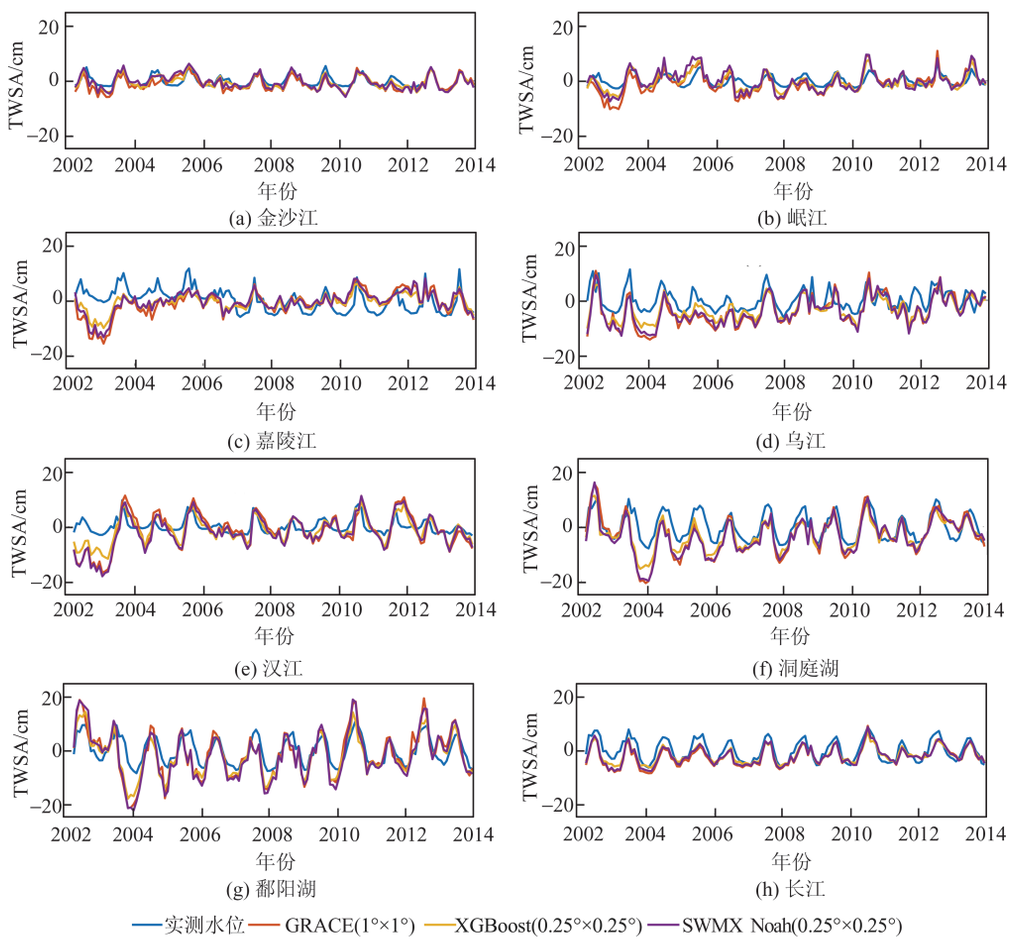


图 9　长江流域及其 7 大支流降尺度前后的 TWSA 与实测水位的时间序列图

Fig. 9　Time Series of TWSA and Measured Water Levels in the Yangtze River Basin and Its Seven Sub-basins Before and After Downscaling

同降尺度方法的差异,还能够评价降尺度后的结果是否优于降尺度前,增益计算式为[47]:

$$G=\frac{|M_o-M_L|-|M_o-M_H|}{|M_o-M_L|+|M_o-M_H|} \tag{9}$$

式中,M_L、M_H分别为降尺度后与降尺度前的度量值;M_o为各度量值的最优值。此处选择R、RMSE和NSE作为度量值。

采用8个水文站实测水位数据来模拟整个流域的时间增益和整个时间序列的空间增益。从图10(a)可以看出,两种SWMX模型和ANN模型结果对应的R、NSE和RMSE的时间增益均存在负值,即降尺度后表现较差。基于SWMX的模型采用CLSM时,降尺度结果的增益略小于采用Noah模型,其原因可能是CLSM TWSA和

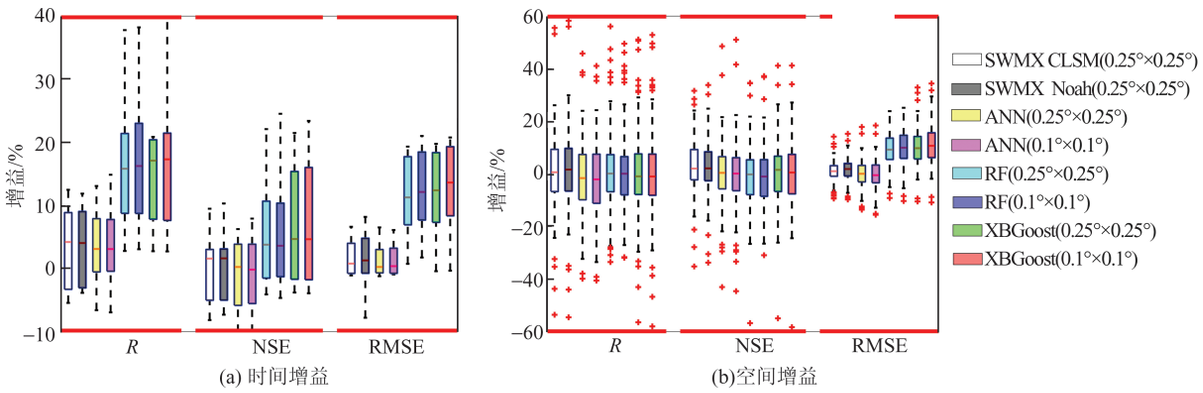


图10 长江流域不同降尺度模型与降尺度至不同分辨率的时间增益和空间增益

Fig. 10 Temporal and Spatial Gains of Different Downscaling Models at Different Spatial Resolutions in the Yangtze River Basin

GRACE TWSA的相关性较Noah低。相比之下,RF和XGBoost方法降尺度后的结果则完全呈现正增益,且降尺度至0.1°×0.1°的结果略优于降尺度至0.25°×0.25°的结果。图10(b)中,降尺度后RMSE的空间增益与各个时间增益结果基本保持一致,而RF和XGBoost模型降尺度结果*R*和NSE的空间增益与其余降尺度模型相比提升幅度较小,表明RF和XGBoost模型降尺度前后的结果在空间尺度上增强幅度不大,这可能与8个水文站数据量太少,计算的*R*和NSE不准确有关。RF和XGBoost模型降尺度的性能较好,且表现基本一致,ANN和SWMX模型表现较差。

3.3.4 降尺度后TWSA空间趋势变化分析

为了更清晰地呈现不同降尺度结果在空间上的差异,图11给出了长江流域2003-02—2017-06降尺度前后TWSA的长期的空间变化特征。图11(a)为降尺度前GRACE TWSA的长期趋势变化,长江上游地区主要呈现减小趋势,中下游地区呈现增加趋势,与现有研究结果一致[48-49]。图11(b)、11(c)分别展示了基于SWMX CLSM和Noah的结果,其长期趋势呈现明显的格网特征,空间分布存在不连续性。图11(d)~11(i)为3种机器学习模型在0.25°×0.25°和0.1°×0.1°分辨率下的长期趋势结果,整体与降尺度前保持较高一致性。其中,RF和XGBoost模型在长江中下游区域(如洞庭湖和鄱阳湖流域)表现出更丰富的细节信息,且两者结果高度一致,与其相近的降尺度性能相符。相比之下,ANN模型降尺度后与原始结果在空间趋势上几乎一致,但这种一致性主要来源于残差增加,而非降尺度效果的提升。

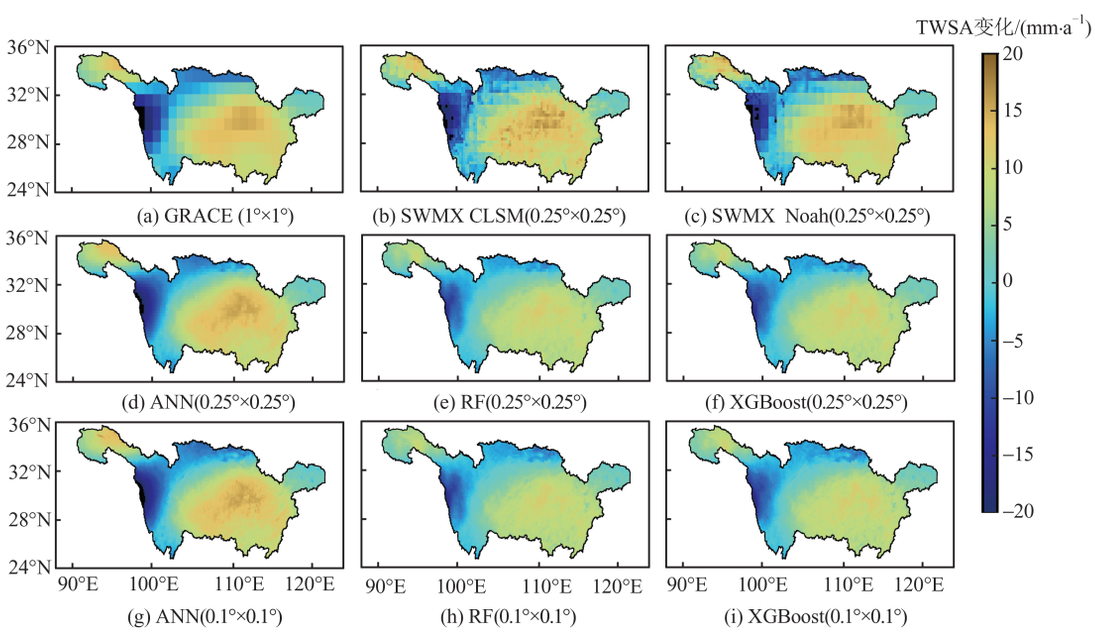


图11 长江流域降尺度前后TWSA长期趋势变化空间分布图

Fig. 11 Spatial Distribution of Long-Term Trends in TWSA Before and After Downscaling in the Yangtze River Basin

以鄱阳湖水体区域为例,图12给出了2002-04—2017-06该区域在降尺度前后长期趋势年变化结果以及相关水文、植被等要素的长期趋势变化。从图12(a)～12(c)可以看出,在1°×1°和0.25°×0.25°下无法区分鄱阳湖水体及陆地区域的TWSA特征。SWMX Noah(0.25°×0.25°)与GRACE TWSA的结果基本一致,而SWMX CLSM(0.25°×0.25°)的长期趋势变化明显变大,与GRACE TWSA结果存在一定的差异,该结果与前文相对增益的评估结果相一致。图12(d)～12(i)中,3种模型结果的长期变化趋势呈现出极大的相似性,尤其是RF和XGBoost模型。图12(j)～12(l)给出了3种VIP得分最高的预测因子的长期趋势变化,将其与机器学习降尺度的3种模型结果进行对比可以看出,3种模型结果均呈现与NDVI和土壤湿度相似的长期趋势变化特征,并且这种变化特征的范围与鄱阳湖水体区域相一致。这表示机器学习模型降尺度在保持原始GRACE数据特征的同时,能够捕捉到植被等要素相关的信号,尤其是PLSR模型重要性高的预测因子。

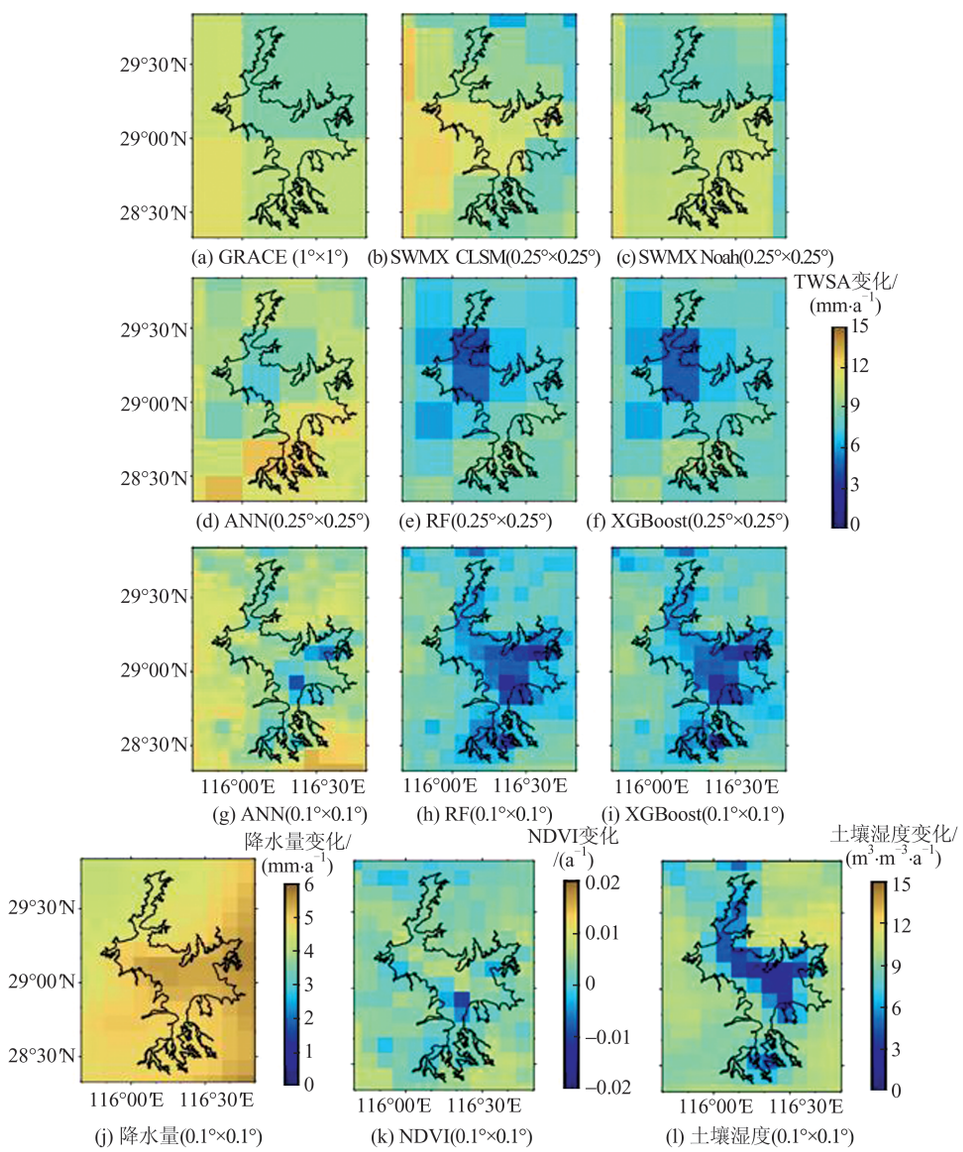


图12　鄱阳湖水体区域降尺度前后TWSA及水文、植被数据的空间趋势变化

Fig. 12　Spatial Trend Variations of TWSA Before and After Downscaling, Along with Hydrological and Vegetation Data, in the Poyang Lake Water Body Region

图13为鄱阳湖水体区域经RF模型降尺度前后TWSA的时间序列变化,降尺度后的数据与实测水位变化更加吻合。表6为多种模型结果与鄱阳湖水位观测数据的R、RMSE和NSE的统计结果,两组SWMX的结果以及ANN模型0.25°×0.25°和0.1°×0.1°的结果相较于GRACE TWSA没有明显改善。RF和XGBoost模型0.25°×0.25°和0.1°×0.1°结果的统计参数均优于降尺度前的结果,同时0.1°×0.1°

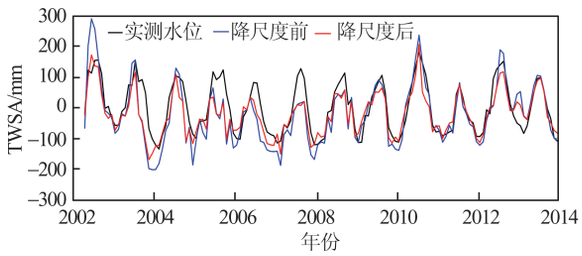


图13 鄱阳湖水体区域RF降尺度前后时间序列变化

Fig. 13 Time Series Changes of RF Downscaling in the Poyang Lake Water Body Region Before and After Downscaling

表6 鄱阳湖水体区域降尺度前后TWSA与水位数据的统计值

Table 6 Statistical Values of TWSA and Water Levels Before and After Downscaling in the Poyang Lake Water Body Region

模型	分辨率	*R*	NSE	RMSE/mm
GRACE	1°×1°	0.840	0.497	58.55
SWMX CLSM	0.25°×0.25°	0.832	0.433	61.26
SWMX Noah	0.25°×0.25°	0.842	0.525	57.14
ANN	0.25°×0.25°	0.845	0.513	57.89
ANN	0.1°×0.1°	0.858	0.541	56.95
RF	0.25°×0.25°	0.870	0.657	50.07
RF	0.1°×0.1°	0.871	0.664	48.32
XGBoost	0.25°×0.25°	0.873	0.673	49.31
XGBoost	0.1°×0.1°	0.881	0.673	48.18

分辨率相较于0.25°×0.25°分辨率结果有一定提升,该结果与长期趋势变化空间分布特征分析结果以及相对增益基本一致。这些研究结果再次证明了RF和XGBoost模型在所有降尺度模型中取得了最佳的降尺度效果。此外,也表明机器学习降尺度模型可以恢复流域水体内更精细的TWSA。

4 结 语

本文首先通过数值模拟方法研究了RF、XGBoost和ANN 3种机器学习降尺度模型用于长江流域TWSA降尺度处理的性能,然后实现了将长江流域GRACE TWSA数据的空间分辨率从1°×1°降尺度到0.25°×0.25°和0.1°×0.1°,并通过与传统水文模型降尺度方法的计算结果和长江流域实测水位数据进行对比和验证,评估了不同机器学习降尺度模型的有效性和差异性。主要研究结论如下:

1)利用PLSR方法计算水文、气象和植被等12个预测因子对长江流域GRACE TWSA的VIP得分,并通过数值模拟实验分析了RF、XGBoost和ANN机器学习模型在不同预测因子下的模型性能和降尺度性能。根据PLSR重要性得分进行预测因子排序及机器学习降尺度结果表明,VIP得分的前6个预测因子(NDVI、土壤湿度、降水、气温、径流和U形风)可以获得较好的效果,增加更多的预测因子对降尺度性能的提升相对较小。

2)通过数值闭合模拟和实测水位数据的对比评估可以看出,传统的水文模型降尺度模型能在一定程度上实现区域分辨率的提高,但水文模型降尺度结果依赖于水文模型与GRACE TWSA数据的相关性。同时,水文模型降尺度模型比机器学习降尺度方法的性能略差,而机器学习方法融合了更多的相关信息,其中RF与XGBoost模型的降尺度性能相当,ANN模型由于模型存在过拟合现象表现略差。

3)3种机器学习降尺度模型的评估结果表明,不同预测因子数量下,不同降尺度方法的相对优劣存在差异。RF和XGBoost模型表现出最佳的降尺度效果,两者的降尺度结果基本一致。机器学习模型中,降尺度到0.1°×0.1°相比0.25°×0.25°的结果略优。此外,对鄱阳湖区域的研究结果表明,RF和XGBoost模型的降尺度结果在保留GRACE主要信号的同时,还能够捕捉到区域更精细的TWSA信息,表明机器学习降尺度模型在提高小尺度的分辨率的同时,能更好地融合水文、气象变化特征,尤其是PLSR中重要性较高的变量。

参 考 文 献

[1] HUMPHREY V, RODELL M, EICKER A. Using Satellite-Based Terrestrial Water Storage Data: A Review[J]. *Surveys in Geophysics*, 2023, 44(5): 1489-1517.

[2] SOLTANI S S, ATAIE-ASHTIANI B, SIMMONS C T. Review of Assimilating GRACE Terrestrial Water Storage Data into Hydrological Models: Advances, Challenges and Opportunities[J]. *Earth-Science Reviews*, 2021, 213: 103487.

[3] 刘任莉,佘敦先,李敏,等. 利用卫星观测数据评估GLDAS与WGHM水文模型的适用性[J]. 武汉大学学报(信息科学版), 2019, 44(11): 1596-1604.
LIU Renli, SHE Dunxian, LI Min, et al. Using Satellite Observations to Assess Applicability of GLDAS and WGHM Hydrological Model[J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(11): 1596-1604.

[4] 钟敏,段建宾,许厚泽,等. 利用卫星重力观测研究近5年中国陆地水量中长空间尺度的变化趋势[J]. 科学通报,2009,54(9):1290-1294.
ZHONG Min, DUAN Jianbin, XU Houze, et al. Using Satellite Gravity Observation to Study the Variation Trend of Land Water Volume in China in the Middle and Long Spatial Scales in Recent Five Years[J]. *Chinese Science Bulletin*, 2009, 54(9): 1290-1294.

[5] 瞿伟,陈沛男,张普方,等. GRACE/GRACE-FO揭示黄河上游2002—2022年GWS及可持续性时空演化[J]. 武汉大学学报(信息科学版),2024,49(8):1287-1299.
QU Wei, CHEN Peinan, ZHANG Pufang, et al. Spatiotemporal Changes of GWS Changes and Their Sustainability in the Upper Reaches of the Yellow River Basin from 2002 to 2022 Revealed by GRACE/GRACE-FO Observations[J]. *Geomatics and Information Science of Wuhan University*, 2024, 49(8): 1287-1299.

[6] CHEN J L, WILSON C R, TAPLEY B D, et al. Long-Term Groundwater Storage Change in Victoria, Australia from Satellite Gravity and In-Situ Observations [J]. *Global and Planetary Change*, 2016, 139: 56-65.

[7] ZHONG B, LI Q, LI X P, et al. Basin-Scale Terrestrial Water Storage Changes Inferred from GRACE-Based Geopotential Differences: A Case Study of the Yangtze River Basin, China[J]. *Geophysical Journal International*, 2023, 233(2): 1318-1338.

[8] 李琼,罗志才,钟波,等. 利用GRACE时变重力场探测2010年中国西南干旱陆地水储量变化[J]. 地球物理学报,2013,56(6):1843-1849.
LI Qiong, LUO Zhicai, ZHONG Bo, et al. Terrestrial Water Storage Changes of the 2010 Southwest China Drought Detected by GRACE Temporal Gravity Field[J]. *Chinese Journal of Geophysics*, 2013, 56(6): 1843-1849.

[9] 肖云,李岩,周泽兵,等. 中国重力卫星进展[J]. 武汉大学学报(信息科学版),2026,51(3):413-424.
XIAO Yun, LI Yan, ZHOU Zebing, et al. Progress on the Chinese Gravimetry Satellite Missions [J]. *Geomatics and Information Science of Wuhan University*, 2026, 51(3): 413-424.

[10] TANGDAMRONGSUB N, DITMAR P G, STEELE-DUNNE S C, et al. Assessing Total Water Storage and Identifying Flood Events over Tonlé Sap Basin in Cambodia Using GRACE and MODIS Satellite Observations Combined with Hydrological Models[J]. *Remote Sensing of Environment*, 2016, 181: 162-173.

[11] CHAO N F, WAN X W, ZHONG Y L, et al. Reconstructing a New Terrestrial Water Storage Deficit Index to Detect and Quantify Drought in the Yangtze River Basin[J]. *Journal of Hydrology*, 2023, 625: 129972.

[12] YI S, SUN W K. Evaluation of Glacier Changes in High-Mountain Asia Based on 10 Year GRACE RL05 Models[J]. *Journal of Geophysical Research: Solid Earth*, 2014, 119(3): 2504-2517.

[13] XIE Y, WANG L S, BEVIS M, et al. Joint Inversion of GNSS and GRACE Data for Ice Mass Loads in Greenland[J]. *Earth and Planetary Science Letters*, 2025, 658: 119329.

[14] 高春春,陆洋,史红岭,等. 基于GRACE RL06数据监测和分析南极冰盖27个流域质量变化[J]. 地球物理学报,2019,62(3):864-882.
GAO Chunchun, LU Yang, SHI Hongling, et al. Detection and Analysis of Ice Sheet Mass Changes over 27 Antarctic Drainage Systems from GRACE RL06 Data [J]. *Chinese Journal of Geophysics*, 2019, 62(3): 864-882.

[15] LIU B S, ZOU X C, YI S, et al. Reconstructing GRACE-Like Time Series of High Mountain Glacier Mass Anomalies[J]. *Remote Sensing of Environment*, 2022, 280: 113177.

[16] ZHANG Z Z, CHAO B F, CHEN J L, et al. Terrestrial Water Storage Anomalies of Yangtze River Basin Droughts Observed by GRACE and Connections with ENSO [J]. *Global and Planetary Change*, 2015, 126: 35-45.

[17] CHEN J L, TAPLEY B, RODELL M, et al. Basin-Scale River Runoff Estimation from GRACE Gravity Satellites, Climate Models, and in Situ Observations: A Case Study in the Amazon Basin[J]. *Water Resources Research*, 2020, 56(10): e2020WR028032.

[18] HUANG Z Y, YEH P J F, JIAO J J, et al. A New Approach for Assessing Groundwater Recharge by Combining GRACE and Baseflow with Case Studies in Karst Areas of Southwest China[J]. *Water Resources Research*, 2023, 59(2): e2022WR032091.

[19] 熊景华,郭生练,王俊,等. 长江流域陆地水储量变化及归因研究[J]. 武汉大学学报(信息科学版),2024,49(12):2241-2248.
XIONG Jinghua, GUO Shenglian, WANG Jun, et al. Variation and Attribution of Terrestrial Water Storage in the Yangtze River Basin[J]. *Geomatics and Information Science of Wuhan University*,

2024, 49(12): 2241-2248.

[20] VON STORCH H, ZORITA E, CUBASCH U. Downscaling of Global Climate Change Estimates to Regional Scales: An Application to Iberian Rainfall in Wintertime[J]. *Journal of Climate*, 1993, 6(6): 1161-1171.

[21] LI B L, RODELL M, ZAITCHIK B F, et al. Assimilation of GRACE Terrestrial Water Storage into a Land Surface Model: Evaluation and Potential Value for Drought Monitoring in Western and Central Europe [J]. *Journal of Hydrology*, 2012, 446: 103-115.

[22] TEUTSCHBEIN C, WETTERHALL F, SEIBERT J. Evaluation of Different Downscaling Techniques for Hydrological Climate-Change Impact Studies at the Catchment Scale[J]. *Climate Dynamics*, 2011, 37(9): 2087-2105.

[23] YIN W J, ZHANG G Q, LIU F T, et al. Improving the Spatial Resolution of GRACE-Based Groundwater Storage Estimates Using a Machine Learning Algorithm and Hydrological Model[J]. *Hydrogeology Journal*, 2022, 30(3): 947-963.

[24] WAN Z M, ZHANG K, XUE X W, et al. Water Balance-Based Actual Evapotranspiration Reconstruction from Ground and Satellite Observations over the Conterminous United States[J]. *Water Resources Research*, 2015, 51(8): 6485-6499.

[25] YIN W J, HU L T, ZHANG M L, et al. Statistical Downscaling of GRACE-Derived Groundwater Storage Using ET Data in the North China Plain [J]. *Journal of Geophysical Research: Atmospheres*, 2018, 123(11): 5973-5987.

[26] ZHANG D, LIU X M, BAI P. Assessment of Hydrological Drought and Its Recovery Time for Eight Tributaries of the Yangtze River (China) Based on Downscaled GRACE Data[J]. *Journal of Hydrology*, 2019, 568: 592-603.

[27] SEYOUM W M, MILEWSKI A M. Improved Methods for Estimating Local Terrestrial Water Dynamics from GRACE in the Northern High Plains [J]. *Advances in Water Resources*, 2017, 110: 279-290.

[28] MIRO M E, FAMIGLIETTI J S. Downscaling GRACE Remote Sensing Datasets to High-Resolution Groundwater Storage Change Maps of California's Central Valley[J]. *Remote Sensing*, 2018, 10(1): 143.

[29] CHEN L, HE Q S, LIU K, et al. Downscaling of GRACE-Derived Groundwater Storage Based on the Random Forest Model [J]. *Remote Sensing*, 2019, 11(24): 2979.

[30] RAHAMAN M M, THAKUR B, KALRA A, et al. Estimating High-Resolution Groundwater Storage from GRACE: A Random Forest Approach[J]. *Environments*, 2019, 6(6): 63.

[31] ZHANG J X, LIU K, WANG M. Downscaling Groundwater Storage Data in China to a 1-km Resolution Using Machine Learning Methods [J]. *Remote Sensing*, 2021, 13(3): 523.

[32] ALI S, KHORRAMI B, JEHANZAIB M, et al. Spatial Downscaling of GRACE Data Based on XGBoost Model for Improved Understanding of Hydrological Droughts in the Indus Basin Irrigation System (IBIS)[J]. *Remote Sensing*, 2023, 15(4): 873.

[33] JANGA B, ASAMANI G P, SUN Z H, et al. A Review of Practical AI for Remote Sensing in Earth Sciences [J]. *Remote Sensing*, 2023, 15 (16): 4112.

[34] KARPATNE A, EBERT-UPHOFF I, RAVELA S, et al. Machine Learning for the Geosciences: Challenges and Opportunities [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2019, 31(8): 1544-1554.

[35] CHENG M K, RIES J C, TAPLEY B D. Variations of the Earth's Figure Axis from Satellite Laser Ranging and GRACE [J]. *Journal of Geophysical Research: Solid Earth*, 2011, 116(B1): B01409.

[36] SWENSON S, WAHR J. Post-Processing Removal of Correlated Errors in GRACE Data[J]. *Geophysical Research Letters*, 2006, 33(8): 2005GL025285.

[37] CHEN J L, WILSON C R, LI J, et al. Reducing Leakage Error in GRACE-Observed Long-Term Ice Mass Change: A Case Study in West Antarctica [J]. *Journal of Geodesy*, 2015, 89(9): 925-940.

[38] WAHR J, MOLENAAR M, BRYAN F. Time Variability of the Earth's Gravity Field: Hydrological and Oceanic Effects and Their Possible Detection Using GRACE [J]. *Journal of Geophysical Research: Solid Earth*, 1998, 103 (B12): 30205-30229.

[39] WANG Y M, LI C, CUI Y J, et al. Spatial Downscaling of GRACE-Derived Groundwater Storage Changes Across Diverse Climates and Human Interventions with Random Forests[J]. *Journal of Hydrology*, 2024, 640: 131708.

[40] CUI L L, LI Y, ZHONG B, et al. Assessing the Impact of 2022 Extreme Drought on the Yangtze River Basin Using Downscaled GRACE/GRACE-FO Data Obtained by Partitioned Random Forest Algorithm[J]. *International Journal of Remote Sensing*, 2025, 46(3): 1219-1247.

[41] AGHBASHLO M, PENG W X, TABATABAEI M, et al. Machine Learning Technology in Biodiesel Research: A Review[J]. *Progress in Energy and Combustion Science*, 2021, 85: 100904.

[42] OTCHERE D A, ARBI GANAT T O, GHOLAMI R, et al. Application of Supervised Machine Learning Paradigms in the Prediction of Petroleum Reservoir Properties: Comparative Analysis of ANN and SVM Models[J]. *Journal of Petroleum Science and Engineering*, 2021, 200: 108182.

[43] DARST B F, MALECKI K C, ENGELMAN C D. Using Recursive Feature Elimination in Random Forest to Account for Correlated Variables in High Dimensional Data[J]. *BMC Genetics*, 2018, 19(1): 65.

[44] XU T R, GUO Z X, LIU S M, et al. Evaluating Different Machine Learning Methods for Upscaling Evapotranspiration from Flux Towers to the Regional Scale[J]. *Journal of Geophysical Research: Atmospheres*, 2018, 123(16): 8674-8690.

[45] CHEN T Q, GUESTRIN C. XGBoost: A Scalable Tree Boosting System[C]//The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, USA, 2016.

[46] WEI A H, YU K N, DAI F G, et al. Application of Tree-Based Ensemble Models to Landslide Susceptibility Mapping: A Comparative Study[J]. *Sustainability*, 2022, 14(10): 6330.

[47] PASCAL C, FERRANT S, SELLES A, et al. Evaluating Downscaling Methods of GRACE (Gravity Recovery and Climate Experiment) Data: A Case Study over a Fractured Crystalline Aquifer in Southern India[J]. *Hydrology and Earth System Sciences*, 2022, 26(15): 4169-4186.

[48] JIANG Q N, FERREIRA V G, CHEN J S. Monitoring Groundwater Changes in the Yangtze River Basin Using Satellite and Model Data[J]. *Arabian Journal of Geosciences*, 2016, 9(7): 500.

[49] 禤键豪, 陈智伟, 张兴福, 等. 联合GRACE与GRACE-FO反演2002—2020年长江流域陆地水储量变化[J]. 大地测量与地球动力学, 2021(9): 961-966.
XUAN Jianhao, CHEN Zhiwei, ZHANG Xingfu, et al. Combining GRACE and GRACE-FO to Derive Terrestrial Water Storage Changes in the Yangtze River Basin from 2002 to 2020[J]. *Journal of Geodesy and Geodynamics*, 2021(9): 961-966.