

严清赞^{1,2} 金双根^{1,2,3} 黄为民⁴ 贾燕⁵ 魏思远¹

基于 CyGNSS 数据的土壤水分与植被光学厚度反演研究

摘要

本文提出了一种仅基于 CyGNSS 数据,能够同时反演土壤水分与植被光学厚度的方案,该方案使用了神经网络与暴力穷举算法.首先考察了 2018 年以及 2020 年的数据,并对结果进行了验证.通过分析发现反演结果与参考数据展现了良好的一致性.土壤水分的反演结果与 2018 年和 2020 年的测试数据比较,其相关系数分别高达 0.86 和 0.84,均方根误差分别为 0.064 和 0.071 cm^3/cm^3 ;对于植被光学厚度,2018 年与 2020 年的相关系数均为 0.98,均方根误差分别为 0.079 和 0.084.研究结果表明,CyGNSS 可作为一种新型且独立的泛热带土壤水分与植被光学厚度反演手段.

关键词

全球导航卫星系统反射测量;土壤水分;神经网络;植被光学厚度

中图分类号 S152.7;S127

文献标志码 A

收稿日期 2021-01-15

资助项目 国家自然科学基金(42001362,42001375)

作者简介

严清赞,男,博士,副教授,研究方向为 GNSS-R 遥感技术与应用.003257@nuist.edu.cn

0 引言

深入探索水文和碳循环是对生物地球物理学的特征要素(例如土壤水分、地表生物多样性等)进行研究的重点^[1-2].卫星遥感因其能有效地提供大空间尺度的数据,成为解决此类问题的一种有效方案.

由于土壤水分(Soil Moisture, SM)影响介电常数,并且微波波段信号对其敏感^[3],因此可以利用微波遥感技术获取大范围的 SM 数据.土壤水分主动-被动探测卫星(SMAP)^[4]和土壤水分和海洋盐度探测卫星(SMOS)^[5]分别搭载了 L 波段辐射计,能够提供基于亮度温度数据的 SM 估计,空间分辨率约为 40 km,每 2~3 天实现全球覆盖.L 波段信号对地表 SM 敏感,且受地表粗糙度和植被的影响较小,因此适用于 SM 的遥感^[6].除了被动式微波遥感接收器,主动遥感系统,例如 Sentinel-1^[7]和 ERRASAR-X^[8]也能够提供 SM 的测量.植被光学深度(VOD,或 τ)作为生物多样性的近似量,可以通过 L 波段 SMAP 和 SMOS 任务以及 C 波段平台,例如高级微波辐射扫描计(AMSR-E)和高级甚高分辨率辐射计(AVHRR)获得^[9].

近几十年来,全球导航卫星系统反射测量(GNSS-R)技术作为一种有效的遥感工具得到了广泛的应用.由于 GNSS-R 搭载的 L 波段信号源能够轻易穿透云层和大气,所以能够实现全时及全天候的遥感监测^[10-11].GNSS-R 已经成功应用于海面风监测^[12-14]、高程测量^[15-16]、海冰遥感^[17-19]、湿地分类^[20]、海啸预警^[21]等方面.随着海量 CyGNSS 数据的公开,利用 CyGNSS 数据估算大尺度的 SM 的研究已成为学者们关注的热点^[22-31].Clarizia 等^[24]提出了一种反射率-植被-粗糙度的算法,利用线性回归函数实现 SM 的反演.Yan 等^[28]采用了类似方法,但计算了 CyGNSS 的观测变量来解析表面粗糙度效应.文献[32-34]研究和评估了从 CyGNSS 数据提取植被相关参数的可行性.

值得注意的是,目前基于 CyGNSS 的研究,无论是估计 SM 还是推导 τ ,通常都依赖于其他辅助数据,例如表面粗糙度、海拔高度等.本文提出了一种通过采用神经网络(Neural Network, NN)算法和暴力穷举法,利用 CyGNSS 产品进行独立的 SM/ τ 反演的有效方案.基于此方案,后续的 SM 和 τ 数据能够同时并直接地从 CyGNSS 数据估计出,且不依赖任何其他的辅助数据.本项工作推进了 CyGNSS 作为 SM 和 τ

1 南京信息工程大学 遥感与测绘工程学院,南京,210044

2 江苏省协同精密导航定位与智能应用工程研究中心,南京,210044

3 中国科学院上海天文台,上海,200030

4 纽芬兰纪念大学(加拿大)应用科学与工程学院,圣约翰斯, A1B3X5

5 南京邮电大学 地理与生物信息学院,南京,210023

数据的独立来源的研究,并向公众打开了一扇 CyGNSS 数据应用新的窗口。

1 数据描述

本节首先描述了 CyGNSS 数据获取与预处理过程,随后介绍了使用的土壤水分 SM 与植被光学厚度 VOD 参考数据及来源。

1.1 CyGNSS 数据

本文使用了 CyGNSS 2.1 版本的一级数据(参见 <https://podaac-tools.jpl.nasa.gov/drive/files/allData/cygnss/L1/v2.1>)。CyGNSS 星座包含了 8 颗微小卫星,能提供覆盖泛热带($38^{\circ}\text{S} \sim 38^{\circ}\text{N}$)的 GNSS-R 数据,该数据有高时空分辨率的特点。本文使用了 2018 年整年以及 2020 年内有可用参考数据的时间段。CyGNSS 一级数据包含了有地理位置信息的时延多普勒图(DDM)、雷达散射截面(BRCS 或 σ)以及其他测量信息与地理坐标信息,如,入射角 θ 、SNR(信噪比)、镜面反射点(SP)的经纬度(lat/lon)以及 SP 到发射机与接收机的距离(R_t 与 R_r)等。本文使用了收集于陆地上且 SNR 大于 0 的 CyGNSS 数据。为了抑制数据中 SP 估计值的误差,本文保留了 BRCS 峰值在时延轴第 4 与 15 位的数据,类似操作也在文献[26,28]中采用。另外, θ 大于 60° 的数据由于质量得不到保证而被剔除[25]。

1.2 参考数据

后续基于 CyGNSS 数据获得的 SM 和 τ 结果会与 SMAP 第 3 级数据进行对比。SMAP 的 SM 数据可以从 <https://nsidc.org/data/SPL3SMP/versions/6> 获取。该数据的空间分辨率为 36 km,大致每 3 天能实现全球覆盖。SMAP 数据提供了 SM、VOD 数据以及质量标记等数据。本文使用了 2018 年全年以及 2020 年第 240 至 291 天质量标记为“反演成功”的数据。为了方便后续的结果验证与比对,对 CyGNSS 数据与 SMAP 产品都进行了 3 d 平均的操作,这样能有效地提高空间覆盖率。另外,本文将 CyGNSS 数据重新处理并投射到 SMAP 数据采用的 EASE-Grid 网格中。

2 估计方法

本节详细叙述了从 CyGNSS 数据独立反演 SM 和 τ 的过程,包括计算 CyGNSS 观测变量、神经网络的构建,以及通过暴力穷举法进行结果计算 3 个步骤。

2.1 CyGNSS 观测量计算

通过假设植被覆盖的光滑面的反射信号是相干的,CyGNSS 的地表反射率 Γ_{RL} 可由以下公式表示[26,35],即:

$$\Gamma_{\text{RL}}(\epsilon_s, \theta) = \mathcal{R}_{\text{RL}}(\epsilon_s, \theta) \gamma^2 \exp(-4k^2 s^2 \cos^2 \theta), \quad (1)$$

式中: θ 是入射角; ϵ_s 表示土壤的介电常数,通常取决于 SM[3];透射率 γ 包含了信号由植被的衰减,通常是 VOD 的函数;指数项代表了表面粗糙度效应, k 为波数, s 是表面的均方根高度;表面的菲涅尔反射系数 \mathcal{R}_{RL} 在 CyGNSS 情况下可以表示为水平极化与垂直极化反射系数分量(\mathcal{R}_{hh} 与 \mathcal{R}_{vv})的线性叠加[36]:

$$\mathcal{R}_{\text{RL}}(\epsilon_s, \theta) = \frac{1}{2} [\mathcal{R}_{\text{vv}}(\epsilon_s, \theta) + \mathcal{R}_{\text{hh}}(\epsilon_s, \theta)], \quad (2)$$

$$\mathcal{R}_{\text{vv}}(\epsilon_s, \theta) = \frac{\epsilon_s \cos \theta - \sqrt{\epsilon_s - \sin^2 \theta}}{\epsilon_s \cos \theta + \sqrt{\epsilon_s - \sin^2 \theta}}, \quad (3)$$

$$\mathcal{R}_{\text{hh}}(\epsilon_s, \theta) = \frac{\cos \theta - \sqrt{\epsilon_s - \sin^2 \theta}}{\cos \theta + \sqrt{\epsilon_s - \sin^2 \theta}}. \quad (4)$$

因此,在植被覆盖的平整表面上的 CyGNSS Γ_{RL} 可认为是 SM、 θ 、VOD 以及表面粗糙度效应的函数。实际中,基于相干反射的假设,CyGNSS Γ_{RL} 可直接从 CyGNSS 的 BRCS(即 σ) 获取[22,24,28,37]:

$$\Gamma_{\text{RL}} = \frac{\sigma(R_t + R_r)^2}{4\pi(R_t R_r)^2}. \quad (5)$$

θ 、 σ 、 R_t 与 R_r 均可从 CyGNSS 数据获取。本文使用相干性标识(CF)来表征表面粗糙度效应,CF 描述了 CyGNSS 数据的功率发散程度[38]。除此之外,经纬度信息(lat/lon)也被用来作为输入[32]。将 CyGNSS Γ_{RL} 、 θ 、CF 以及 lat/lon 都基于 EASE-Grid 网格取逐 3 d 平均。作为例子,图 1 展示了这 5 项以及 SMAP 的 SM 与 VOD 在 2018 年的平均值。

2.2 神经网络的构建

正如前文所提,本文假设了 Γ_{RL} 、 θ 、CF(ϵ_{CF})、lat/lon(ϵ_{lat} 、 ϵ_{lon})、SM(ζ_{SM})以及 VOD(τ)之间的相互作用。自然而然地,它们之间的关系可以表达如下:

$$\zeta_{\text{SM}} = f(\Gamma_{\text{RL}}, \theta, \epsilon_{\text{CF}}, \epsilon_{\text{lat}}, \epsilon_{\text{lon}}, \tau), \quad (6)$$

$$\tau = g(\Gamma_{\text{RL}}, \theta, \epsilon_{\text{CF}}, \epsilon_{\text{lat}}, \epsilon_{\text{lon}}, \zeta_{\text{SM}}). \quad (7)$$

获取这两个公式严格的数学表达式有较大难度。这里,借助机器学习算法来构建函数 f 与 g 。神经网络算法已成功地在 GNSS-R 遥感领域证明了其有效性与准确性,例如海冰监测[39]、风速反演[14]以及土壤/植被遥感[26-27,32]。本文也采用神经网络算法。

考虑有 n 个样本的训练集 $T = \{(\mathbf{x}_1, y_1), \dots,$

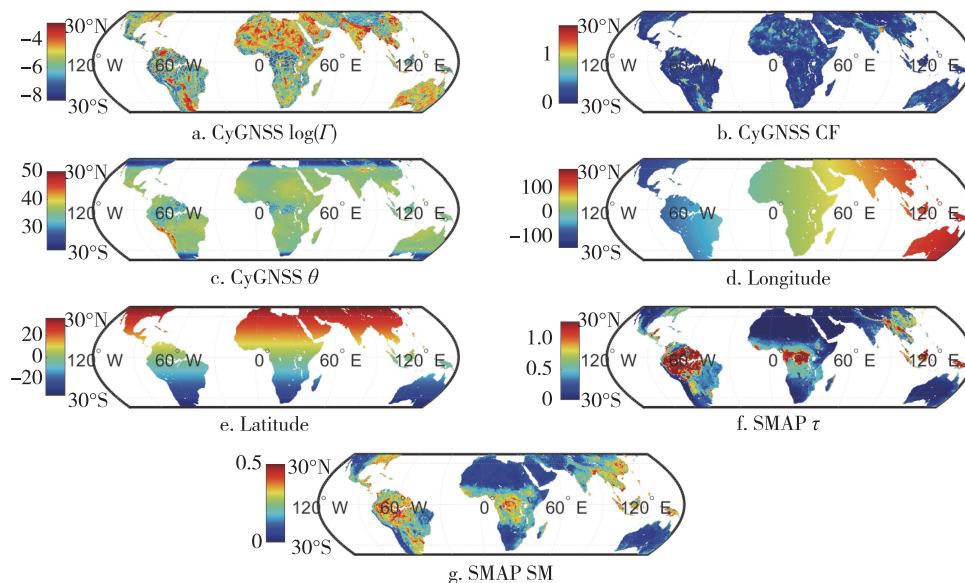


图 1 2018 年 CYGNSS 与 SMAP 数据的平均值

 Fig. 1 Annual mean of: (a) CyGNSS Γ_{RL} , (b) CF, (c) θ , (d) lon, (e) lat, (f) SMAP τ , and (g) SMAP SM in 2018

(\mathbf{x}_n, y_n) , 其中 \mathbf{x}_j 是输入向量 (包含 Γ_{RL} 、 θ 、CF、lat/lon 与 τ , 且 $j = 1, \dots, n$), 而 y_j 代表了式 (6) 中的 SM 值. 在这里, 神经网络算法可以看作用来构建函数 f , 从而实现 $f(x)$ 高精度地近似 SM (或 y). 同理, 为了估算 VOD, 可以参照式 (7). 图 2 给出了神经网络的结构示意, 由 1 个输入层、3 个隐藏层以及 1 个输出层构成. 每一层的神经元个数可由使用者任意设定, 相邻层的神经元都由权重相连接, 除输入层外神经元的激励过程总结如下:

$$\mathbf{a}^{(k+1)} = \phi(\mathbf{W}^{(k+1)} \mathbf{a}^{(k)} + \mathbf{b}^{(k+1)}), k \in [0, 1, 2, 3], \quad (8)$$

式中 $\mathbf{a}^{(k+1)}$ 、 $\mathbf{b}^{(k+1)}$ 与 $\mathbf{W}^{(k+1)}$ 分别是第 $(k+1)$ 层的激

励向量、偏差向量以及权重矩阵. ϕ 是激励函数. 另一方面, 输入层的表达式如下:

$$\mathbf{a}^{(0)} = \mathbf{x}. \quad (9)$$

式 (8) 与 (9) 总结了神经网络的前向传播路径. 一旦网络完成学习过程, 即权重与偏差都通过训练而确定, 便可以通过式 (6) 与 (7) 的形式来估计 SM 与 VOD. 在网络训练阶段, 采用后向传播学习^[40] 以及 LM 算法^[41], 它们的本质在于, 在给定训练数据集 T 时, 通过调整权重与偏差值将以下的误差函数最小化:

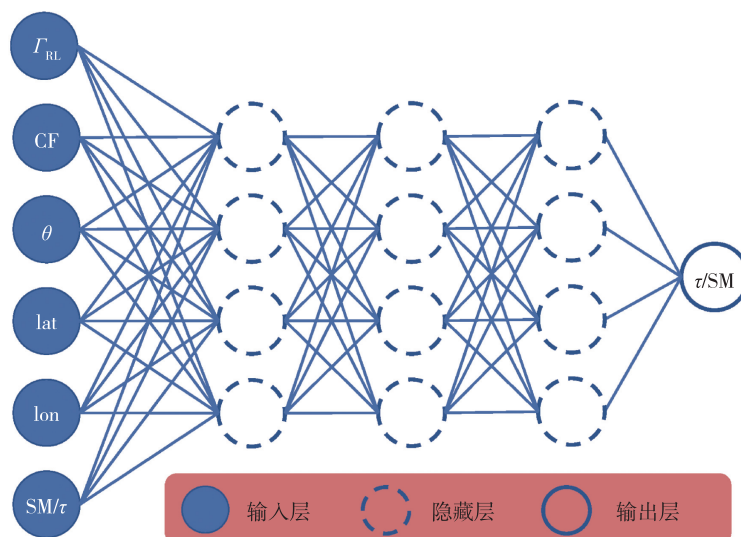


图 2 神经网络结构示意

Fig. 2 An example of an NN

$$\varepsilon(W, b; T) = \frac{1}{2} \sum_{i=1}^n \|\hat{y}_i - y_i\|^2, \quad (10)$$

式中 W 与 b 分别代表权重与偏差值的集合, \hat{y}_i 是估计值。

总的来说,本文使用了神经网络算法来构建 f 与 g , 进而通过式(6)与(7)的形式实现对 SM 与 VOD 的估计.详细的神经网络算法使用过程可参照文献[14,39].本文对神经网络的构建、训练以及后续分析都基于 MATLAB R2019b 软件。

2.3 独立反演方案

当 f 与 g 构建完成后,从后续的 CyGNSS 数据反演 SM/VOD 即可通过暴力穷举法实现,其详细步骤如下.首先,将式(6)中的 τ 替换成式(7),可以得到:

$$\zeta_{SM} = f[\Gamma_{RL}, \theta, \varepsilon_{CF}, \varepsilon_{lat}, \varepsilon_{lon}, g(\Gamma_{RL}, \theta, \varepsilon_{CF}, \varepsilon_{lat}, \varepsilon_{lon}, \zeta_{SM})], \quad (11)$$

或以一种更简洁的形式,即:

$$\zeta_{SM} = F(\Gamma_{RL}, \theta, \varepsilon_{CF}, \varepsilon_{lat}, \varepsilon_{lon}, \zeta_{SM}), \quad (12)$$

或者写成:

$$\zeta_{SM} - F(\Gamma_{RL}, \theta, \varepsilon_{CF}, \varepsilon_{lat}, \varepsilon_{lon}, \zeta_{SM}) = 0. \quad (13)$$

至此,对于 VOD 的依赖便消除了.理想情况下,只有准确的 SM 与其对应的 CyGNSS 数据才能使式(13)成立.在实际操作中,我们使用暴力穷举法来求取 SM 的估计值.将不同的 SM 备选值作为输入,并记录式(13)等式左侧的差值,把对应于最小绝对值

的 SM 作为估计.同时,确定 SM 值后, VOD 可通过式(7)计算出来.如图 3 所示,该流程总结了本文所提出的独立反演方案及其具体步骤。

3 结果与评价

本节对提出的 SM 和 τ 反演方案进行了测试,并且对 2018 与 2020 年的数据进行了评估.根据 SMAP 数据的可用性,本文选取了 2018 全年以及 2020 年第 240 天至 291 天的数据.基于第 2 节描述的数据匹配方案,分别获得了 2018 年 1 620 多万和 2020 年 230 多万的数据样本。

3.1 神经网络的训练

随机选择了 2018 年数据中的 1% 进行神经网络的训练,即以 f 的形式(式(6))从 Γ_{RL} 、 θ 、 ε_{CF} 、 ε_{lat} 、 ε_{lon} 以及 VOD 获取 SM,以 g 的形式(式(7))获取 VOD.另一方面,剩下的数据,包括 2018 年剩余的 99% 和 2020 年的数据,都将用作最终的反演验证。

神经网络结构的设计(如层数或神经元的数目,以及激励函数的选择)可以是任意的,并可能导致网络的精确度不同.本文测试了大量不同结构的神经网络,来得到“最优”的网络,将隐藏层数设置为 1、2、3,神经元数为 1~50,并将所有被测网络的激励函数设置为“tansig”^[29],其形式如下:

$$\phi(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)}. \quad (14)$$

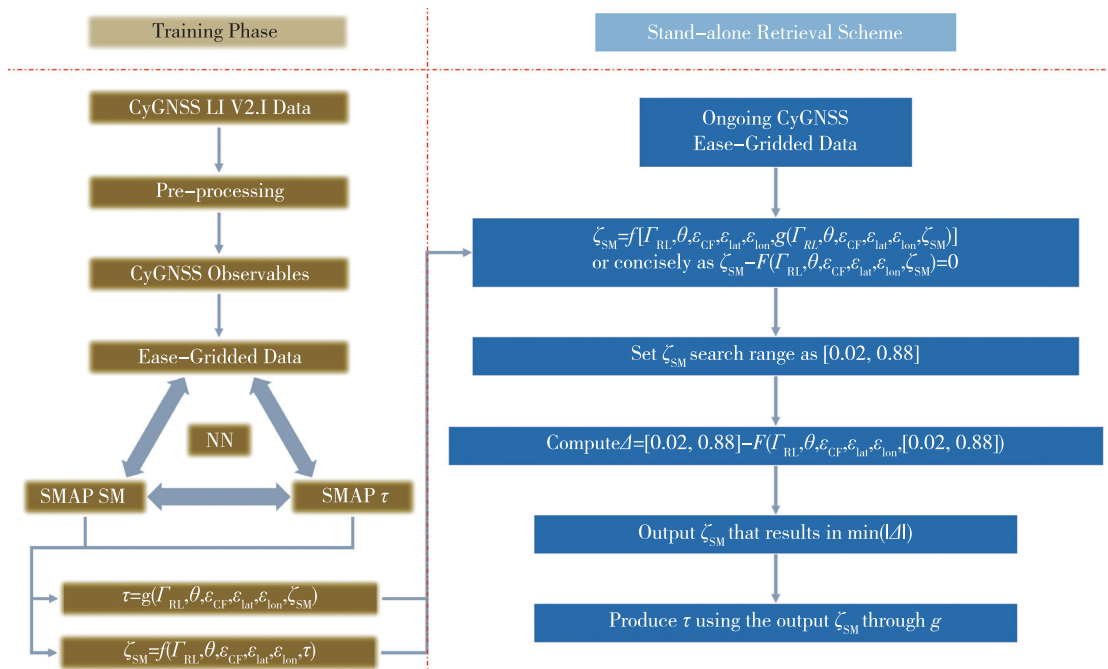


图 3 独立反演 SM 与 VOD 的方案流程

Fig. 3 A flowchart presents the steps of data preparation, networking training and the final retrieval

通过对测试的神经网络进行准确度评价,最终选择了有 3 个隐藏层,每个隐藏层有 26 个神经元组成的网络来进行对 SM 的估计,即 f ;另外,有 3 个隐藏层,每个隐藏层有 38 个神经元构成的网络被用作 τ 的预测器.图 4 中显示了这 2 种网络性能的密度.目前为止,只有 1% 的 2018 年数据运用到了神经网络的训练中.

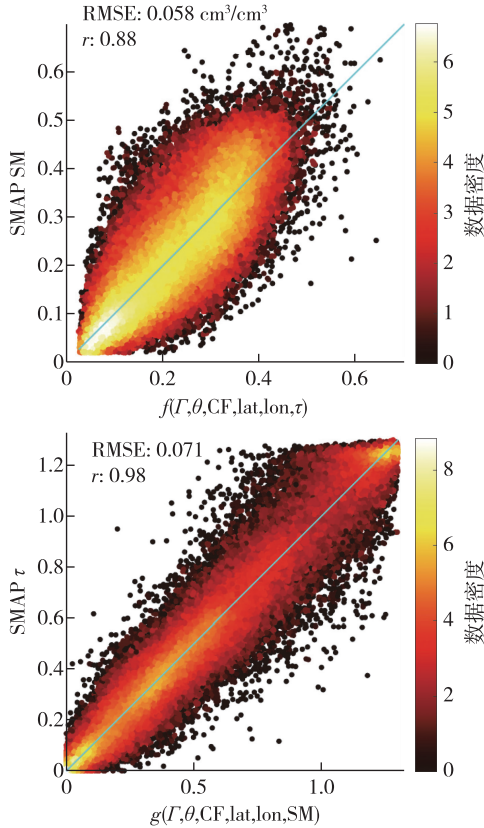


图 4 数据密度图与性能

Fig. 4 Density plot showing the performance of employed NN

3.2 SM 与 VOD 的反演

确定好神经网络 f 和 g , 按照 2.3 节的步骤对 SM 和 τ 进行反演.根据训练数据的范围和精度,设置 SM 的搜索范围为 $0.02 \sim 0.88 \text{ cm}^3/\text{cm}^3$,步长为 $10^{-4} \text{ cm}^3/\text{cm}^3$.值得注意的是,用于反演 SM/ τ 的输入样本只需要 CyGNSS 的数据,包括 I_{RL} 、 θ 、CF、lat 与 lon.除去之前用于神经网络训练的数据之外,2018 年和 2020 年剩余的所有测试数据都用于反演结果的评估.在实际操作中,SM 是根据计算出的差值 $\Delta = (0.02:10^{-4}:0.88) - F[I_{\text{RL}}, \theta, \varepsilon_{\text{CF}}, \varepsilon_{\text{lat}}, \varepsilon_{\text{lon}}, (0.02:10^{-4}:0.88)]$ 进行评估的,结果中对应 $\min(|\Delta|)$ 的值被视作 SM. τ 则通过式 (7) 计算得出.当 Δ 中多次出现交叉零点时(说明有多解),则认为反演无效.这

可能是由网络固有误差引起的.尽管如此,仅约 6% 的数据因为这个问题而被舍弃.

上述反演过程基于式 (11) — (13) 的推导.具体来说,是将变量 τ 替换了式 (6) 中的 g .在现实中,这个过程可以采用另一种方式,即将式 (7) 中的 f 由 SM 代替,那么可得到与式 (13) 对等的形式:

$$\tau - G(I_{\text{RL}}, \theta, \varepsilon_{\text{CF}}, \varepsilon_{\text{lat}}, \varepsilon_{\text{lon}}, \tau) = 0. \quad (15)$$

显然,也可以使用式 (15) 和上述的类似方法,对 SM 和 τ 进行新一轮的反演.为提高结果的可靠性,本文利用式 (15) 也进行了反演,为 τ 设置的搜索范围为 $[0, 1.2970]$,步长为 10^{-4} .只有当基于式 (13) 和 (15) 得到的反演结果一致时,才认为最终的反演结果是有效的.为了检验并量化结果一致性的影响,我们测试了用以筛选反演结果的不同阈值,范围为 $[10^{-5}, 10^{-1}]$.如图 5 所示,当两种方案的反演结果吻合度越高,所得到的最终结果与参考数据也更一致.考虑到所获得的准确结果,本文将采用 10^{-5} 作为阈值进行更深层次的评价和讨论.

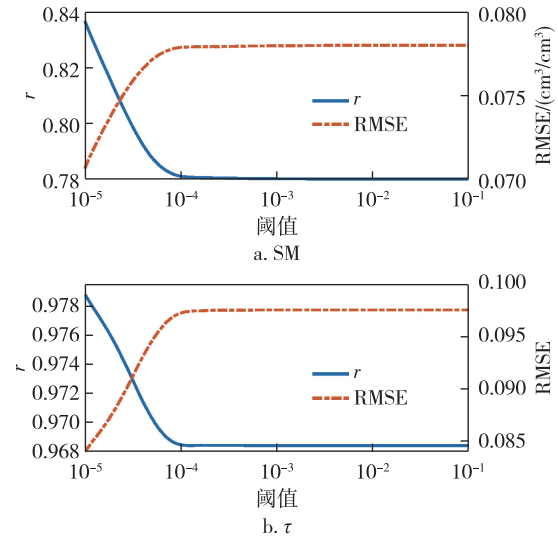


图 5 不同反演方案所得估计值间的一致性对最终结果的影响

Fig. 5 Impact of result agreement on final retrieval accuracy:
(a) SM, and (b) τ estimations

3.3 验证与评估

如图 6 所示,基于训练数据与测试数据的反演结果均展现了高准确度.相关系数 r 和均方根误差 RMSE 汇总在表 1 中.结果表明,本文所提出的反演方案具有较好的通用性和可靠性.特别地,尽管测试数据集(2020 年)与训练数据集(2018 年)存在一定的时间差距,但仍能利用 2020 年数据成功地进行反

演,验证了该方案的有效性.并且,训练数据和测试数据的密度之间的高度一致性也是显而易见的.

由于2018年只有1%的数据参与了神经网络的训练,所以测试集的结果相对于训练集的结果有一个不显著下降是合理且可接受的.2018年和2020年数据的反演性能的差异可能是由于总体的数据质量存在着潜在差异.值得注意的是,2020年CyGNSS数据的非相干求和时间(0.5 s)比2018年的(1 s)短,

因此前者的空间采样率是后者的两倍.这对整体的数据质量带来了两种相反的影响.一方面,SNR与非相干求和时间的平方根成正比^[42],这使得2018年CyGNSS的个体数据都具有更高的SNR;另一方面,2020年的数据具有更高的空间采样率,这将导致在每个Ease-Grid网格中有更多的样本,从而提高了精度^[28].因此,2018年和2020年网格化后的数据质量可能存在一定的差异.

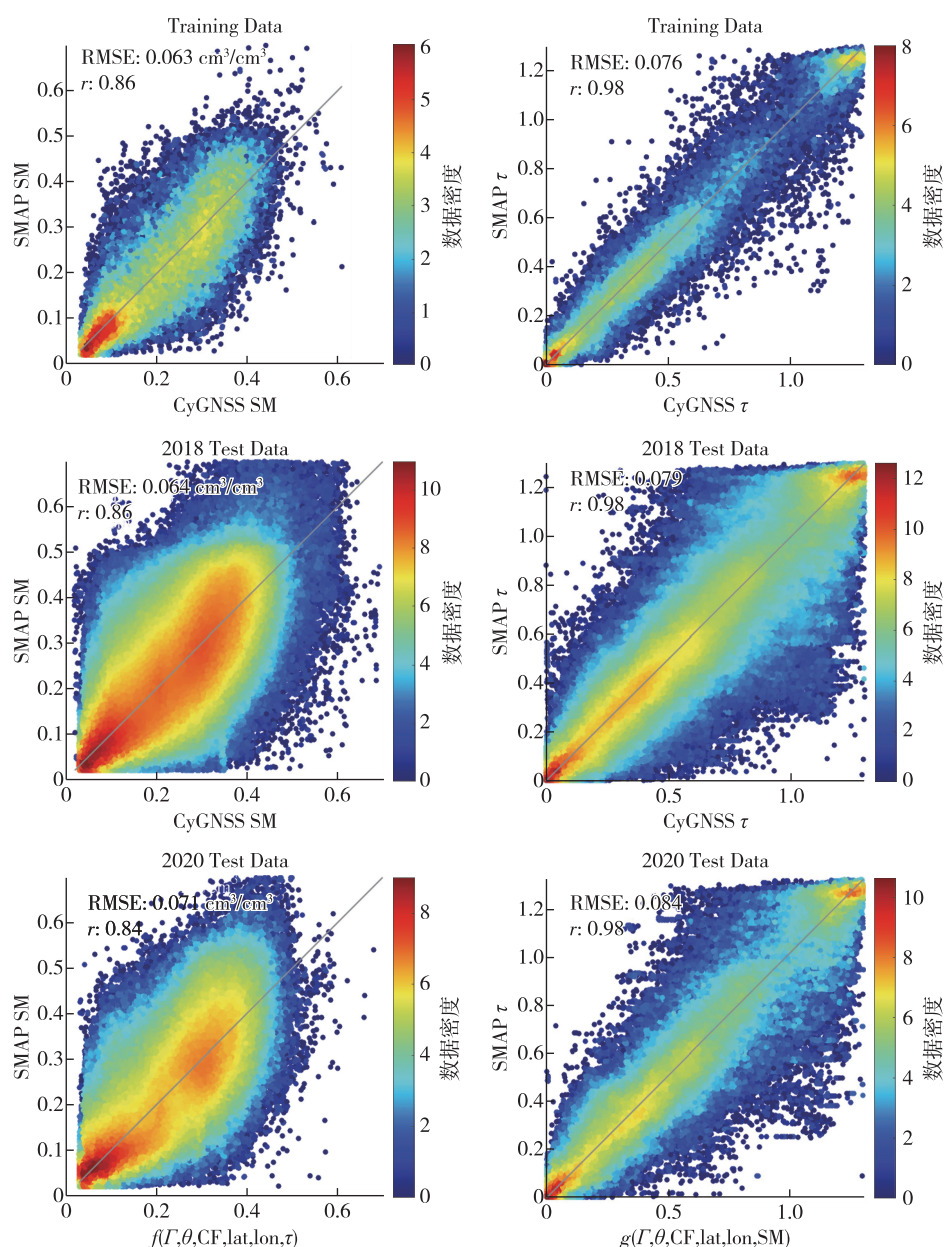


图6 最终SM与VOD反演结果的密度

Fig. 6 Density plots of final SM and VOD retrievals using training and test data

表 1 最终反演的性能统计
Table 1 Performance statistics of final retrievals

反演模型	VOD				SM			
	<i>r</i>		RMSE		<i>r</i>		RMSE/(cm ³ /cm ³)	
	<i>g</i>	<i>F & G</i>	<i>g</i>	<i>F & G</i>	<i>f</i>	<i>F & G</i>	<i>f</i>	<i>F & G</i>
训练数据	0.99	0.98	0.067	0.076	0.89	0.86	0.056	0.063
测试数据 (2018)	0.98	0.98	0.071	0.079	0.89	0.86	0.058	0.064
测试数据 (2020)	0.98	0.98	0.076	0.084	0.87	0.84	0.063	0.071

下面进一步验证 SMAP 和 CyGNSS 数据产品时空变化响应的一致性.因为 2018 年的数据量更大,这里展示了基于 2018 年的数据结果.如图 7 所示,以年平均的形式呈现了 SMAP 和 CyGNSS 结果之间的空间相关性(SM 见图 7a 和 7b, τ 见图 8a 及 8b),并且通过比对 SM 和 τ 时间序列的标准差(Std),证实了时间尺度上的一致性(图 7、图 8c、8d).SMAP 与 CyGNSS 数据结果间的差异很小.

4 结束语

本文提出了一种基于 CyGNSS 数据同时反演土壤水分和植被光学厚度的方法.这种方法结合了神经网络技术与暴力穷举算法.设计的输入包括由 CyGNSS 导出的反射率、相干性标识、入射角以及经纬度.两种不同的变量替换方式形成了两种不同的

反演模型(具体形式见式(13)与(15)).为了增强该方案的可靠性,只有当两种模型的反演结果一致时,我们才认为反演结果是有效的,并评估了其对结果准确性的影响.通过最终反演结果与 SMAP 数据的对比验证,发现反演结果与参考数据展现了良好的一致性,验证了该反演模型的有效性和准确性.针对 2018 年与 2020 年的土壤水分比对结果,得到的相关系数分别达 0.86 与 0.84,均方根误差分别为 0.064 和 0.071 cm³/cm³.对于植被光学厚度,2018 年与 2020 年的结果均得到 0.98 的相关系数,均方根误差分别为 0.079 和 0.084.本文进一步从时空变化的角度验证了反演结果的可靠性.今后,在有同时同地参考数据的情况下,将使用统一非相干求和时间的 CyGNSS 数据以便消除整体数据质量的不确定性.如何确保神经网络系统的效率、适用性和精确性是至

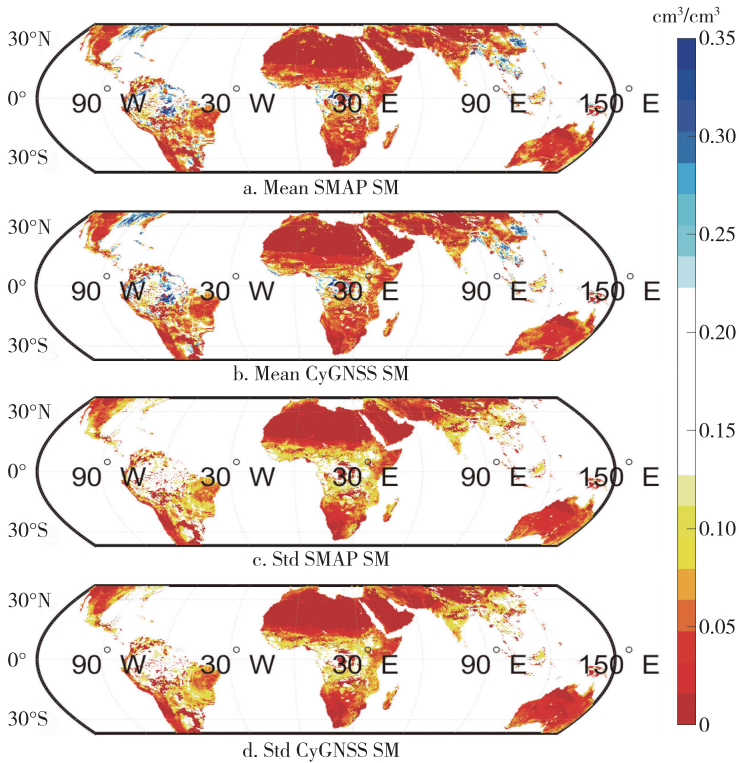


图 7 SM 结果一致性比对
Fig. 7 Distributions of SM results:(a) mean SMAP,(b) mean CyGNSS,(c) Std SMAP,and (d) Std CyGNSS

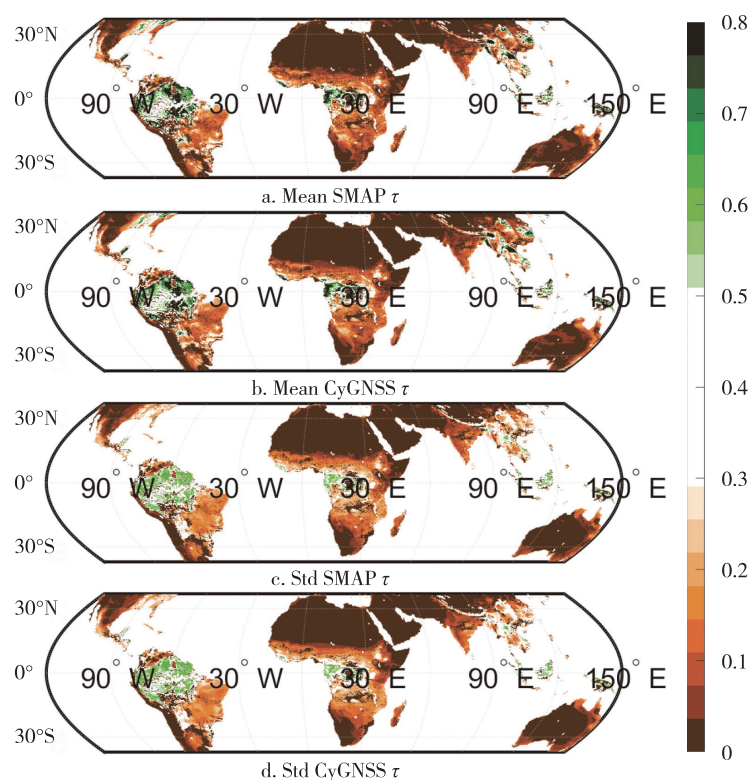


图8 VOD结果一致性比对

Fig. 8 Patterns of τ results:(a) mean SMAP,(b) mean CyGNSS,(c) Std SMAP, and (d) Std CyGNSS

关重要的,值得进一步研究和改进.此外,本文神经网络是基于 SMAP 数据进行调参的,SMAP 数据也可以被其他具有更高分辨率的数据或现场测量数据所替代.并且,本文采用暴力穷举法获得的最终反演结果,其初始值的设置与步长精度与最终性能高度相关.因此,扩大搜索范围和细化步长可以进一步改善结果.暴力穷举法的替代方法也将在未来研究中进一步探索.

参考文献

References

- [1] Dobriyal P, Qureshi A, Badola R, et al. A review of the methods available for estimating soil moisture and its implications for water resource management[J]. Journal of Hydrology, 2012, 458/459: 11-117
- [2] Bennett A C, Penman T D, Arndt S K, et al. Climate more important than soils for predicting forest biomass at the continental scale [J]. Ecography, 2020, 43 (11): 1692-1705
- [3] Dobson M C, Ulaby F T, Hallikainen M T, et al. Microwave dielectric behavior of wet soil-part II: dielectric mixing models[J]. IEEE Transactions on Geoscience and Remote Sensing, 1985, GE-23(1):35-46
- [4] Entekhabi D, Njoku E G, O'Neill P E, et al. The Soil Moisture Active Passive (SMAP) mission [C] // Proceedings of the IEEE, 2010, 98(5):704-716
- [5] Kerr Y H, Waldteufel P, Wigneron J P, et al. Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission[J]. IEEE Transactions on Geoscience and Remote Sensing, 2001, 39 (8): 1729-1735
- [6] Saillard M, de Roo R D, Ulaby F T. Comments on "Bistatic specular scattering from rough dielectric surfaces" [with reply] [J]. IEEE Transactions on Antennas and Propagation, 1995, 43(2):224-226
- [7] 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
- [8] Aubert M, Baghdadi N, Zribi M, et al. Analysis of TerraSAR-X data sensitivity to bare soil moisture, roughness, composition and soil crust[J]. Remote Sensing of Environment, 2011, 115(8):1801-1810
- [9] Vittucci C, Vaglio Laurin G, Tramontana G, et al. Vegetation optical depth at L-band and above ground biomass in the tropical range: evaluating their relationships at continental and regional scales [J]. International Journal of Applied Earth Observation and Geoinformation, 2019, 77:151-161
- [10] Jin S G, Komjathy A. GNSS reflectometry and remote sensing: new objectives and results [J]. Advances in Space Research, 2010, 46(2):111-117
- [11] Jin S G, Feng G P, Gleason S. Remote sensing using GNSS signals: current status and future directions [J].

- Advances in Space Research, 2011, 47(10):1645-1653
- [12] Foti G, Gommenginger C, Jales P, et al. Spaceborne GNSS reflectometry for ocean winds: first results from the UK TechDemoSat-1 mission [J]. Geophysical Research Letters, 2015, 42(13):5435-5441
 - [13] Clarizia M P, Ruf C S, Jales P, et al. Spaceborne GNSS-R minimum variance wind speed estimator[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(11):6829-6843
 - [14] Liu Y X, Collett I, Morton Y J. Application of neural network to GNSS-R wind speed retrieval[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(12):9756-9766
 - [15] Cardellach E, Rius A, Martín-Neira M, et al. Consolidating the precision of interferometric GNSS-R ocean altimetry using airborne experimental data [J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(8):4992-5004
 - [16] Li W Q, Cardellach E, Fabra F, et al. Assessment of spaceborne GNSS-R ocean altimetry performance using CYGNSS mission raw data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(1):238-250
 - [17] Yan Q Y, Huang W M. Spaceborne GNSS-R sea ice detection using delay-Doppler maps: first results from the U. K. TechDemoSat-1 mission [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9(10):4795-4801
 - [18] Rodriguez-Alvarez N, Holt B, Jaruwatanadilok S, et al. An Arctic sea ice multi-step classification based on GNSS-R data from the TDS-1 mission[J]. Remote Sensing of Environment, 2019, 230:111202
 - [19] Yan Q Y, Huang W M. Sea ice thickness measurement using spaceborne GNSS-R: first results with TechDemoSat-1 data[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13:577-587
 - [20] Rodriguez-Alvarez N, Podest E, Jensen K, et al. Classifying inundation in a tropical wetlands complex with GNSS-R[J]. Remote Sensing, 2019, 11(9):1053
 - [21] Yan Q Y, Huang W M. Tsunami detection and parameter estimation from GNSS-R delay-Doppler map [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9(10):4650-4659
 - [22] Chew C C, Small E E. Soil moisture sensing using spaceborne GNSS reflections: comparison of CYGNSS reflectivity to SMAP soil moisture [J]. Geophysical Research Letters, 2018, 45(9):4049-4057
 - [23] Kim H, Lakshmi V. Use of Cyclone Global Navigation Satellite System (CyGNSS) observations for estimation of soil moisture[J]. Geophysical Research Letters, 2018, 45(16):8272-8282
 - [24] Clarizia M P, Pierdicca N, Costantini F, et al. Analysis of CYGNSS data for soil moisture retrieval[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(7):2227-2235
 - [25] Al-Khaldi M M, Johnson J T, O'Brien A J, et al. Time-series retrieval of soil moisture using CYGNSS[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(7):4322-4331
 - [26] Eroglu O, Kurum M, Boyd D, et al. High spatio-temporal resolution CYGNSS soil moisture estimates using artificial neural networks [J]. Remote Sensing, 2019, 11(19):2272
 - [27] Senyurek V, Lei F N, Boyd D, et al. Machine learning-based CYGNSS soil moisture estimates over ISMN sites in CONUS[J]. Remote Sensing, 2020, 12(7):1168
 - [28] Yan Q Y, Huang W M, Jin S G, et al. Pan-tropical soil moisture mapping based on a three-layer model from CYGNSS GNSS-R data[J]. Remote Sensing of Environment, 2020, 247:111944
 - [29] Santi E, Pettinato S, Paloscia S, et al. Soil moisture and forest biomass retrieval on a global scale by using CyGNSS data and artificial neural networks[C]// International Geoscience and Remote Sensing Symposium, 2020:5905-5908
 - [30] Jia Y, Jin S G, Savi P, et al. Modeling and theoretical analysis of GNSS-R soil moisture retrieval based on the random forest and support vector machine learning approach[J]. Remote Sensing, 2020, 12(22):3679
 - [31] Yan Q Y, Gong S Q, Jin S G, et al. Near real-time soil moisture in China retrieved from CyGNSS reflectivity[J]. IEEE Geoscience and Remote Sensing Letters, 2020, 99:1-5
 - [32] Santi E, Paloscia S, Pettinato S, et al. Remote sensing of forest biomass using GNSS reflectometry [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13:2351-2368
 - [33] Carreno-Luengo H, Luzi G, Crosetto M. Sensitivity of CyGNSS bistatic reflectivity and SMAP microwave radiometry brightness temperature to geophysical parameters over land surfaces[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(1):107-122
 - [34] Carreno-Luengo H, Luzi G, Crosetto M. Above-ground biomass retrieval over tropical forests: a novel GNSS-R approach with CyGNSS [J]. Remote Sensing, 2020, 12(9):1368
 - [35] Choudhury B J, Schmugge T J, Chang A, et al. Effect of surface roughness on the microwave emission from soils [J]. Journal of Geophysical Research: Oceans, 1979, 84(C9):5699-5706
 - [36] Zavorotny V U, Voronovich A G. Scattering of GPS signals from the ocean with wind remote sensing application[J]. IEEE Transactions on Geoscience and Remote Sensing, 2000, 38(2):951-964
 - [37] Carreno-Luengo H, Camps A, Querol J, et al. First results of a GNSS-R experiment from a stratospheric balloon over boreal forests[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(5):2652-2663
 - [38] Al-Khaldi M M, Johnson J T, Gleason S, et al. An algorithm for detecting coherence in cyclone global navigation satellite system mission level-1 delay-Doppler maps[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 99:1-10
 - [39] Yan Q Y, Huang W M, Moloney C. Neural networks based sea ice detection and concentration retrieval from GNSS-

- R delay-Doppler maps [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(8): 3789-3798
- [40] Werbos P. Beyond regression: new tools for prediction and analysis in the behavioral sciences [D]. Cambridge: Harvard University, 1974
- [41] Marquardt D W. An algorithm for least-squares estimation of nonlinear parameters [J]. Journal of the Society for Industrial and Applied Mathematics, 1963, 11(2): 431-441
- [42] Gleason S. Space-based GNSS scatterometry: ocean wind sensing using an empirically calibrated model [J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(9): 4853-4863

Retrievals of soil moisture and vegetation optical depth using CyGNSS data

YAN Qingyun^{1,2} JIN Shuanggen^{1,2,3} HUANG Weimin⁴ JIA Yan⁵ WEI Siyuan¹

1 School of Remote Sensing & Geomatics Engineering, Nanjing University of Information Science & Technology, Nanjing 210044

2 Jiangsu Engineering Center for Collaborative Navigation/Positioning and Smart Applications, Nanjing University of Information Science & Technology, Nanjing 210044

3 Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030

4 Faculty of Engineering and Applied Science, Memorial University, St. John's, NL A1B3X5

5 School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing 210023

Abstract In this paper, a new scheme is proposed for simultaneously retrieving Soil Moisture (SM) and Vegetation Optical Depth (VOD), solely from the Cyclone Global Navigation Satellite System (CyGNSS) data. This work is accomplished by employing two pre-trained neural networks (NNs), including one for computing SM from the CyGNSS data and Soil Moisture Active Passive (SMAP) VOD as well as the other for calculating VOD from the CyGNSS data and SMAP SM product, through a brute-force searching. By adopting the proposed method, the posterior SM/VOD can be estimated merely using the CyGNSS data, free from other auxiliary data. The attained results are validated against SMAP products for two separate periods: the whole year of 2018 and a recent course in 2020. Satisfactory agreements between the retrieved and referred SM/VOD are achieved, with correlation coefficients (r) of 0.86 and 0.84, along with root-mean-square errors (RMSEs) of 0.064 and 0.071 cm^3/cm^3 for SM in the years of 2018 and 2020, respectively; and for the verification of VOD, $r=0.98$ and $\text{RMSE}=0.079$ are acquired for 2018, and $r=0.98$ and $\text{RMSE}=0.084$ for 2020, respectively. The good consistency obtained in this work illustrates the capability of CyGNSS as a new independent source for estimating pan-tropical SM and VOD.

Key words Global Navigation Satellite System-Reflectometry (GNSS-R); Soil Moisture (SM); Neural Network (NN); Vegetation Optical Depth (VOD)