IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10642855

DISENTANGLING VEGETATION WATER CONTENT AND SURFACE SOIL MOISTURE FROM SENTINEL-1 SAR TIME SERIES THROUGH ASYNCHRONOUS CONVOLUTIONAL NEURAL NETWORK

Changjiang Shi[1,2,3], Zhijie Zhang[4,], Baohui Wang[5], Wanchang Zhang[1,2] and Yaning Yi[6]*

[1] Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
[2] International Research Center of Big Data for Sustainable Development Goals, Beijing 100094, China
[3] University of Chinese Academy of Sciences, Beijing 100049, China
[4] Department of Environment and Society, Quinney College of Natural Resources, Utah State University, Logan, UT 84322, USA
[5] School of Software, Beihang University, Beijing 100191, China
[6] National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing 100085, China
*Corresponding author: zhijiezhang@arizona.edu

ABSTRACT

This paper introduces and assesses an algorithm tailored for the discrimination of Vegetation Water Content (VWC) and Surface Soil Moisture (SM) using time series data acquired from the Sentinel-1 Synthetic Aperture Radar (SAR). The backscatter signals captured from the Earth's surface encompass contributions from both soil surface scattering and multiple interactions involving both the soil surface and canopy constituents. The conventional water cloud model encounters challenges in soil moisture retrieval due to the intricate interplay of moisture in both vegetation and soil, rendering direct retrieval from the backscatter coefficient a complex task. Employing the principle of blind signal separation, we deploy an asynchronous convolutional neural network to distinctively retrieve vegetation water and soil moisture from the Sentinel-1 SAR VV polarization backscatter time series. Our findings validated at OzNet network reveal an unbiased Root Mean Square Error (RMSE) of 0.054 m^3/m^3 for soil moisture retrieval, with the accuracy of vegetation water content closely aligning with SMAP auxiliary VWC data.

Index Terms— Soil Moisture Retrieval, Vegetation Water Content Retrieval, Signal Separation, SAR Time Serial, Deep Learning

1. INTRODUCTION

Accurate soil moisture data, characterized by fine spatial resolution (less than 1 km) and frequent temporal updates (2–3 days), are vital across various sectors. Current global remote sensing soil moisture (SM) products, derived from space-borne scatterometers and radiometers, offer spatial resolutions in the order of 10s of kilometers. However, attempts to enhance spatial resolution through downscaling algorithms have fallen short, especially in achieving kilometer-level resolution. Recent advancements in space-borne Synthetic Aperture Radars (SAR), coupled with improved revisit frequency and an emerging open data policy, present a promising alternative for obtaining high-resolution SM data. Utilizing data from the Sentinel-1 constellation, pre-operational 1-km SM products are now accessible at regional to continental scales, leveraging time-series data and assuming time-invariant roughness and vegetation, categorized into three main groups.

Multi-temporal scattering model inversion methods extend snapshot approaches [1] by leveraging temporal information for SM retrieval [2]. This approach significantly improves inversion stability by addressing unknowns such as roughness and vegetation variations, which pose challenges in snapshot methods. To further tackle data and model uncertainties, an ensemble framework was introduced [3], aggregating retrievals from diverse channels and timepoints to reduce overall error. Despite achieving good accuracy, implementing robust scattering models with wide applicability across landcover types and fine resolutions remains challenging for large-scale implementations.

The other two methods involve long-term change detection (LTCD) and short-term change detection (STCD). The LTCD method, introduced by [4], estimates soil wetness by linearly scaling backscatter time series. Conversely, the alpha approximation method [5], categorized as STCD, derives absolute SM by relating the ratio of consecutive backscatter observations to SM variation. Both approaches have been expanded to incorporate multi-polarized data and techniques for bounding radar-derived SM. Building upon these change detection methods, the advanced change detection (ACD) method [6] refines SM retrieval from Sentinel-1 data by: 1) accounting for the influence of temporal vegetation changes through two-way attenuation modeling and 2) incorporating a temporal SM constraint derived from the coarse SMAP

product, thereby mitigating uncertainties caused by vegetation and/or surface roughness variations. The assumption of time-invariant roughness and vegetation is crucial for successful time series SM retrieval. While acceptable in STCD and model-based time series methods, this assumption challenges LTCD, particularly at high resolutions where vegetation dynamics within the retrieval window are significant. Unlike its predecessors, ACD utilizes additional SMAP observations, introducing constraints from vegetation and roughness, albeit at the expense of uncertainties from the SMAP data itself.

This study introduces a novel blind signal separation method based on deep learning to disentangle temporal vegetation evolution and soil moisture from C-band Sentinel-1 data. Leveraging recent advancements in neural network architectures, particularly for sequence modeling [7], this method promises improved performance in separating these intertwined signals. The complex relationship between radar backscatter, SM, and vegetation makes direct SM retrieval challenging. The backscattered signal arises from a combination of surface and volume scattering from soil and vegetation, forming a "mixed" time series. Our deep learning model addresses this challenge by directly processing the mixed backscatter signal to simultaneously retrieve SM and VWC.

2. DATA AND PREPROCESSING

The performance assessment was carried out over the OzNet hydrological monitoring network [8] with 34 stations, and the stations are located in a way that they cover the prevailing environmental conditions in the area according to their respective fractions. We collected 5 cm soil moisture observations from this network between August 21, 2015, and December 31, 2020. 50% of the stations were randomly selected as training samples and the remaining 50% as validation samples.

This study utilized 160 scenes of Sentinel-1A descending Interferometric Wide Ground Range Detected data covered the OzNet soil moisture network acquired between August 21, 2015, and December 31, 2020. The pre-processing steps included calibration, co-registration, geocoding, and temporal filtering of the VV backscatter coefficients, resulting in a stack of VV backscatter coefficients time series data with 100m pixel resolution.

Vegetation water content data from the SMAP L3 ancillary data product as a validation data source for the separation algorithm to extract vegetation water content. Since SMAP and Sentinel-1A transit times are different, but SMAP has a smaller revisit period, the SMAP VWC data was interpolated to obtain the VWC value corresponding to the Sentinel transit time.

3. METHODOLOGY

The proposed separation algorithm in this study endeavors to extract individual time series of VWC and Surface SM from Sentinel-1 VV mixture backscatters time series by end to end neural network model.

3.1 Sources of backscatter

The traditional water cloud model [9] defines backscatter sources in VV-polarization as:

Direct vegetation scattering:

$$\sigma_{\mathrm{vv}}^{v} = AV_1 \cos\,\theta_i (1-\tau^2) \tag{1}$$

$$\tau^2 = \exp\,(-2BV_2 \sec\,\theta_i) \tag{2}$$

Soil-attenuated scattering:

$$\tau^2 \sigma_0^s = \exp\,(-2BV_2 \sec\,\theta_i)\sigma_0^s \tag{3}$$

Direct soil scattering:

$$\sigma_0^s = C + D * SM \tag{4}$$

Neglecting negligible vegetation-soil interaction in co-polarized radiation, the total VV backscatter becomes:

$$\begin{aligned}\sigma_{\mathrm{vv}} = {} & AV_1 \cos\,\theta_{\mathrm{i}}(1-\exp\,(-2BV_2 \sec\,\theta_i) \\ & +\exp\,(-2BV_2 \sec\,\theta_i)(C + D * SM)\end{aligned} \tag{5}$$

whereis the direct scattering from the vegetation canopy (i.e. volume scattering), $\sigma_{\mathrm{vv}}^{s}$is the direct scattering from the soil surface, τ^2 is the two-way vegetation attenuation. θ_i is the incident angle, A and B are the vegetation parameters, C and D are bare soil parameters, V_1 and V_2 are the vegetation descriptors and SM is the volumetric soil moisture. Where A, B, C, D and E are the parameters to be determined during the fitting of the water cloud model.

Since the angle of incidence of the backscatter data collected in the area where the site is located is around 39 degrees, the effect of the angle is not considered in the task of separating vegetation water content and soil moisture. We denote the time series of the mixture backscatters as $\boldsymbol{X} \in \mathbb{R}^{1\times T}$, thus we simplify the VV backscatter time series from Sentinel SAR as follows:

$$\boldsymbol{X} = F_v(\boldsymbol{s}_{vmc}) + F_s(\boldsymbol{s}_{vmc}, \boldsymbol{s}_{sm}) + n \tag{6}$$

where $F_v(\boldsymbol{s}_{vmc})$and $F_s(\boldsymbol{s}_{vmc}, \boldsymbol{s}_{sm}) \in \mathbb{R}^{1\times T}$ denotes the backscatters time series from VWC and SM sources,$\boldsymbol{s}_{vmc}$ and $\boldsymbol{s}_{sm} \in \mathbb{R}^{1\times T}$ denotes VWC and SM time series.$n \in \mathbb{R}^{1\times T}$ denotes the noise signal, T denotes the length of the time series.

3.2 Convolutional neural network for separation

The proposed network architecture, detailed in Figure 1, processes short backscatter segments through three stages: encoding, source separation, and decoding. The encoder

maps segments to intermediate feature representations used to estimate source-specific multiplicative functions per time step. Reconstructed vegetation water content and soil moisture time series are then obtained by applying the decoder to masked encoder features. Each stage is further elaborated in the following subsections.

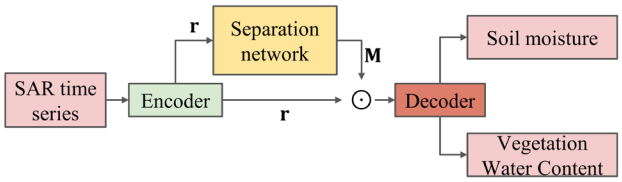

Figure1. Overview of the Separation model

3.2.1 Convolutional Encoder and Decoder

The encoder divides $\boldsymbol{X}$ into K overlapping segments $\bar{\mathbf{x}}_k \in \mathbb{R}^{1\times L}$ and transforms each segment into a feature vector $\bar{\mathbf{r}}_k \in \mathbb{R}^{L\times N}$:

$$\bar{\mathbf{r}}_k = \bar{\mathbf{x}}_k \mathbf{U}_\mathrm{e} \tag{7}$$

The source separation process involves three steps. First, applying a trainable 1D convolution with kernel $\mathbf{U}_\mathrm{e}$ and appropriate stride to $\mathbf{U}_\mathrm{e} \in \mathbb{R}^{L\times N}$, the weight matrix. Then estimating a mask $\mathbf{M}_i \in \mathbb{R}^{1\times N}$ for source i using the separation network (detailed in Section 3.2.2). This network receives $\bar{\mathbf{r}}_k$ and employs a fully connected layer with ReLU activation to generate $\mathbf{M}_i$. Finally, the decoder reconstructs the SM and VWC segment $\bar{\mathbf{s}}_{sm,k}$, $\bar{\mathbf{s}}_{vmc,k}$

$$\bar{\mathbf{s}}_{i,k} = \left(\bar{\mathbf{r}}_k \odot \mathbf{M}_i\right)\mathbf{U}_\mathrm{d}^\top \tag{8}$$

The weight matrix $\mathbf{U}_\mathrm{d} \in \mathbb{R}^{L\times N}$ is applied to the output from the separation network and Encoder through element-wise multiplication ($\odot$). Estimated states $\widehat{\boldsymbol{s}_{vmc}}$ and $\widehat{\boldsymbol{s}_{sm}}$ are then obtained by summing K overlapping segments, $\bar{\mathbf{s}}_{i,k}$, via a 1-D transposed convolution operation.

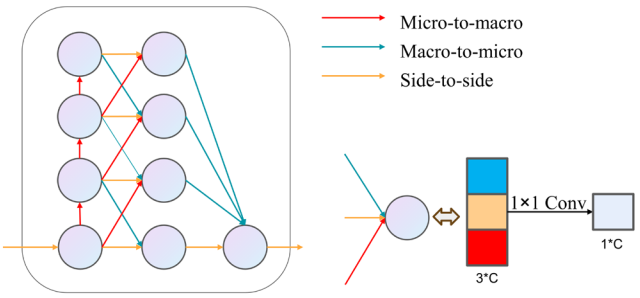

Figure 2. Asynchronous Convolutional Block

3.2.2 Separation network

The Asynchronous Convolutional Block (ACB) [10] serves as the fundamental unit within our separation network. This graph-based block is structured with nodes representing stages (scales) and edges symbolizing information flow (refer to Figure 2, where S=4 denotes the number of stages). Each node corresponds to a convolutional layer processing a different scale, with higher nodes outputting coarser information.

The ACB integrates three connection types:

- Micro-to-macro: Sequential updates within and between stages (Figure 2, illustrated by vertical arrows).
- Macro-to-micro: Information flow from higher to lower stages.
- Side-to-side: Fusion between adjacent stages (Figure 2, indicated by horizontal arrows).

Information processing in the ACB unfolds in three phases:

1) Sequential updates: Processing from the bottom stage to the upper stage.
2) Adjacent stage fusion: Simultaneous stage updates through information exchange.
3) Global fusion: Fusing information from all stages back to the bottom stage.

Each stage accommodates C feature maps. Multi-scale fusion takes place at the stage input by concatenating K input feature maps (resulting in K*C channels) and subsequently reducing them to C channels via a 1x1 convolutional layer.

This multi-stage processing emulates local and global information fusion seen in biological systems. Connections between stages employ the same operation and parameters, reflecting corresponding neuron sets. The overall separation network comprises multiple interconnected ACBs, as depicted in Figure 3, with the number of ACBs set to 8 in this study.

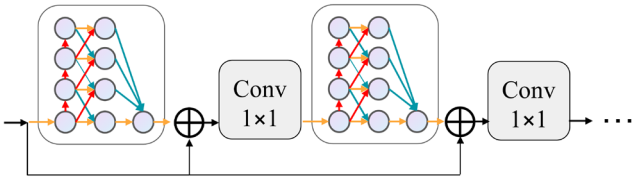

Figure 3. Architecture of the separation network

To train the proposed model to separate vegetation water content and soil surface moisture, a MSE loss function is adopted in this study. For a set of observations $\boldsymbol{s}_{sm}$, $\boldsymbol{s}_{vmc}$ and corresponding model predictions $\widehat{\boldsymbol{s}_{sm}}$, $\widehat{\boldsymbol{s}_{vmc}}$, loss is calculated as follows:

$$Loss = \frac{1}{t}\sum_{i=1}^{T}\left(\boldsymbol{s}_{sm_i} - \widehat{\boldsymbol{s}_{sm,i}}\right)^2 + \left(\boldsymbol{s}_{vmc} - \widehat{\boldsymbol{s}_{vmc,i}}\right)^2 \tag{9}$$

In the 50 epoch training phase, we employed Adam optimizer and a cosine annealing scheduler for learning rate adjustment. The learning rate ranged from 0.001 to 0.0005.

The evaluative framework for the deep learning-based model was appraised through the examination of four key indicators: correlation coefficient (CC), bias, root mean square error (RMSE), and unbiased RMSE (ubRMSE).

4. RESULTS AND DISCUSSION

The performance of the SM and VWC separation algorithm is rigorously evaluated using temporal series from

the OzNet in situ network and SMAP auxiliary product. Figure 4 illustrates the validation station results, showcasing a strong agreement between the model's estimated SM and the in-situ measurements. The correlation coefficient stands at 0.595, with a minimal bias of 0.011 m^3/m^3, an RMSE of 0.055 m^3/m^3, and an unbiased RMSE of 0.054 m^3/m^3. Turning to the retrieved VWC, when compared to SMAP VWC data (refer to Figure 5), the model output demonstrates an impressive correlation coefficient of 0.920. Additionally, there is a negligible bias of 0.081 kg/m^2, an RMSE of 0.271 kg/m^2, and an unbiased RMSE of 0.259 kg/m^2.

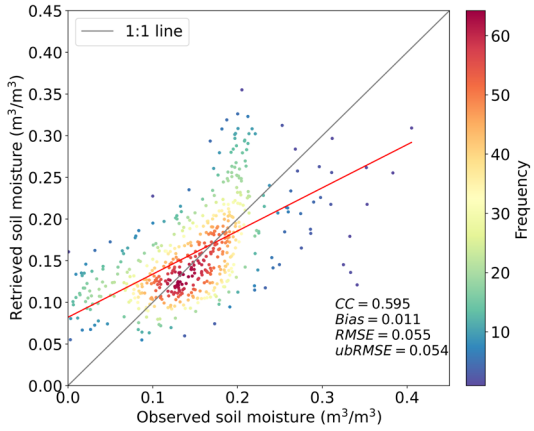

Figure 4. OzNet network intensive in-situ versus retrieved soil moisture

These outcomes affirm the high accuracy achieved by the proposed model in simultaneously retrieving both SM and VWC, particularly noteworthy considering only Sentinel backscatter data were utilized.

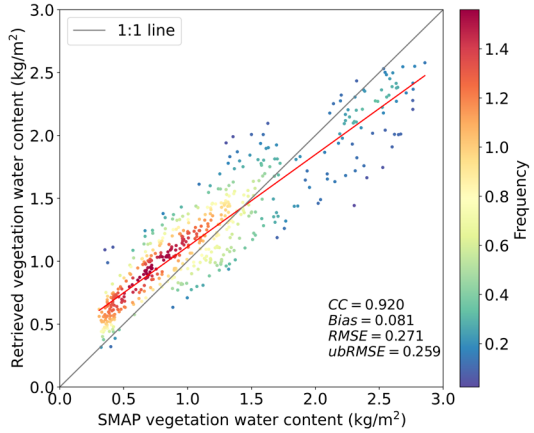

Figure 5. SMAP auxiliary vegetation water content versus retrieved vegetation water content

5. CONCLUSIONS

We introduce a convolutional separation network, a deep learning framework designed for end-to-end SAR signal separation. This model efficiently retrieves soil moisture and vegetation water content simultaneously from Sentinel SAR backscatter time series, employing a time series blind signal separation approach. The performance of the SM and VWC separation algorithm is rigorously evaluated using temporal series from the OzNet in situ network and SMAP auxiliary product. The model's estimations exhibit a commendable level of consistency with the reference data. The ubRMSE between the retrieved SM and the measured sites stands at 0.054 m^3/m^3, while the ubRMSE between the retrieved VWC and the reference SMAP VWC is 0.259 kg/m^2. These results underscore the robustness and accuracy of the proposed model in disentangling SM and VWC from Sentinel-1 SAR time series data. We anticipate further advancements in the model's development, with updates to be presented at IGARSS 2024.

REFERENCES

[1] K. C. Kornelsen and P. Coulibaly, "Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications," *Journal of Hydrology*, vol. 476, pp. 460–489, Jan. 2013, doi: 10.1016/j.jhydrol.2012.10.044.

[2] N. Pierdicca, L. Pulvirenti, and C. Bignami, "Soil moisture estimation over vegetated terrains using multitemporal remote sensing data," *Remote Sensing of Environment*, vol. 114, no. 2, pp. 440–448, Feb. 2010, doi: 10.1016/j.rse.2009.10.001.

[3] L. Zhu, J. P. Walker, and X. Shen, "Stochastic ensemble methods for multi-SAR-mission soil moisture retrieval," *Remote Sensing of Environment*, vol. 251, p. 112099, Dec. 2020, doi: 10.1016/j.rse.2020.112099.

[4] W. Wagner, G. Lemoine, and H. Rott, "A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data," *Remote Sensing of Environment*, vol. 70, no. 2, pp. 191–207, Nov. 1999, doi: 10.1016/S0034-4257(99)00036-X.

[5] A. Balenzano, F. Mattia, G. Satalino, and M. W. J. Davidson, "Dense Temporal Series of C- and L-band SAR Data for Soil Moisture Retrieval Over Agricultural Crops," *IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing*, vol. 4, no. 2, pp. 439–450, Jun. 2011, doi: 10.1109/JSTARS.2010.2052916.

[6] L. Zhu, "Time series soil moisture retrieval from SAR data: Multi-temporal constraints and a global validation," *Remote Sensing of Environment*, 2023.

[7] Z.-Q. Wang and D. Wang, "Combining Spectral and Spatial Features for Deep Learning Based Blind Speaker Separation," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 27, no. 2, pp. 457–468, Feb. 2019, doi: 10.1109/TASLP.2018.2881912.

[8] A. B. Smith *et al.*, "The Murrumbidgee soil moisture monitoring network data set," *Water Resources Research*, vol. 48, no. 7, p. 2012WR011976, Jul. 2012, doi: 10.1029/2012WR011976.

[9] Long, David, and Fawwaz Ulaby. *Microwave radar and radiometric remote sensing*. Artech, 2015.

[10] M. Cannici, M. Ciccone, A. Romanoni, and M. Matteucci, "Asynchronous Convolutional Networks for Object Detection in Neuromorphic Cameras," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Long Beach, CA, USA: IEEE, Jun. 2019, pp. 1656–1665. doi: 10.1109/CVPRW.2019.00209.