

A NEW APPROACH FOR DOWNSCALING SOIL MOISTURE BY MERGING CONDITIONAL ADVERSARIAL NETWORKS AND PHYSICS-BASED PASSIVE MICROWAVE RETRIEVAL

Archana Kannan¹, Grigorios Tsagkatakis², Mahta Moghaddam¹

¹University of Southern California, Los Angeles, CA, United States, ²Foundation for Research and Technology -Hellas (FORTH), Institute of Computer Science, Heraklion, Greece

ABSTRACT

Numerous satellite sources measure global soil moisture, but the spatial resolution of these products is usually coarse making it challenging to use them for various science applications. In this work, we develop a conditional generative adversarial network-based model to downscale satellite measured brightness temperature. Augmenting a physics-based *Tau-omega* model to the framework, high resolution soil moisture is retrieved. For training and validation, upscaled *in-situ* soil moisture measurements are utilized. Experimental results suggests that the proposed framework captures the soil moisture variations from *in-situ* sensors and helps in substantially improving the resolution of passive microwave satellite soil moisture maps.

Index Terms— Deep Learning, cGAN, Downscaling, Soil Moisture, *Tau-omega* model

1. INTRODUCTION

Soil Moisture (SM) is an essential climate variable needed by various hydrological, meteorological, climatological, and environmental applications. In the recent times there is an abundance of satellite measurements from which SM is being estimated using various microwave active and passive retrieval algorithms. Though these satellite images are accessible on global scale, they often come in coarse spatial resolutions (e.g., 36 km and 9 km from Soil Moisture Active Passive (SMAP) satellite observations). But many of the agricultural, wild fire monitoring, weather, and geo-hazard prediction applications require much finer spatial resolution of SM [1] in the 100 m to 1 km scales.

Various approaches have been proposed in literature to obtain fine resolution SM [1] [2]. These approaches include upscaling *in-situ* sensor measurements into spatial resolutions on the order of 100 m. But it can be seen that the value of soil moisture between adjacent observations can have high variability both spatially and temporally [3]. Furthermore, it is simply not feasible to deploy *in-situ* soil moisture networks over extended regions globally. Hence using *in-situ* SM networks is not ideal or practical for obtaining fine resolution SM. Downscaling satellite observations is another ap-

proach for obtaining high resolution SM. Downscaling using statistical methods, geo-information methods, and model-based methods have been extensively reported [2]. Methods that involve usage of multi-source satellite data are especially seen in the recent times for obtaining high resolution SM. SMAP proposed a baseline algorithm that disaggregates soil moisture products from its passive sensor with backscatter data from an active sensor [4]. There are also algorithms that downscale Brightness Temperature (TB) first from a passive sensor and subsequently use inversion techniques to obtain SM [1]. In addition to physics-based fusion approaches, various machine intelligence based methods were also proposed for downscaling SM, where satellite SM images were disaggregated using other ancillary data. Sun and Cui (2021) [5] used a feed forward neural network to downscale SM to 1 km resolution. Convolutional Neural Networks (CNNs) were used to downscale SM to 1 km in [6] and [7].

In this work, we combine both physics-based and machine learning methods to enhance the resolution of SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture [8]. This approach first disaggregates 9 km spatial resolution TB to 1 km TB using a modified U-NET [9] based Convolutional Neural Network (CNN) architecture. U-NET architectures were originally proposed for image segmentation applications. This structure is being applied for downscaling satellite imagery in this work. 1 km spatial resolution SM is retrieved from 1 km spatial resolution TB predictions using the well-known first order radiative transfer model, the *Tau-omega* model [10].

Additionally, to account for the significant differences in resolution scales and the inherent noise of the TB measurement process, another Deep Neural Network (DNN) is augmented with U-NET to form a conditional Generative Adversarial Network (cGAN) [11]. This allows a bit of flexibility from typical rigid loss functions (such as mean absolute errors) and gives a score on whether the predicted image is “realistic”.

The downscaling framework and training inputs are discussed in detail in sections 2 and 3 respectively.

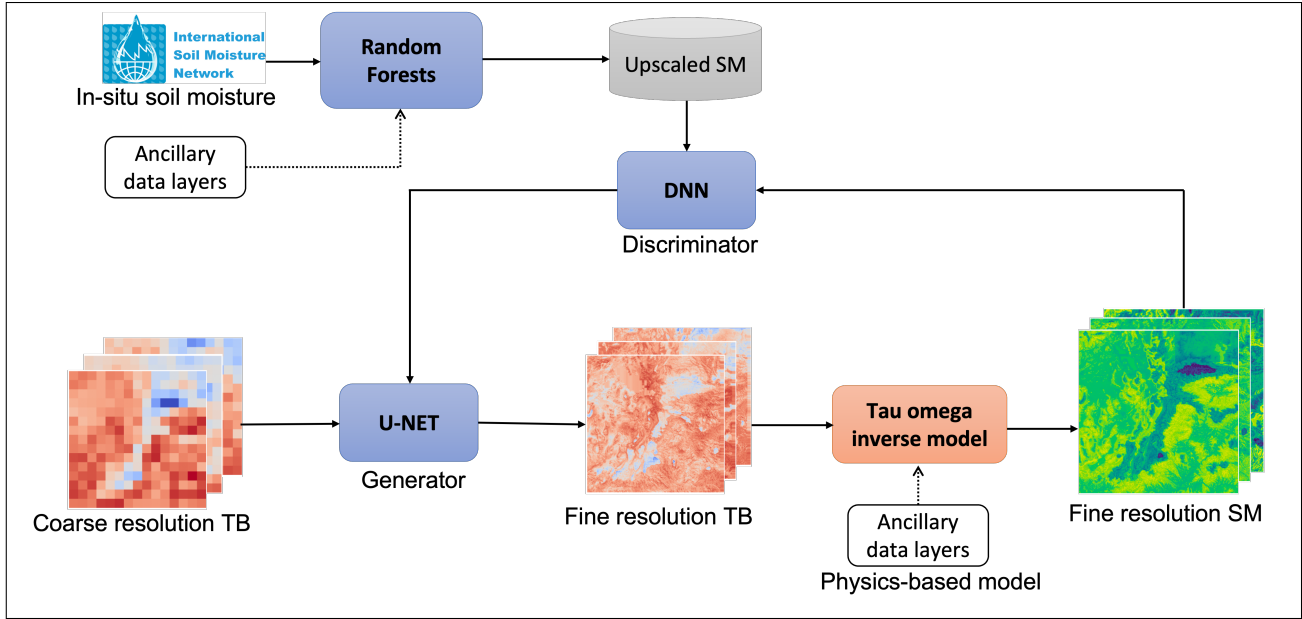


Fig. 1. Proposed cGAN integrative framework for SMAP TB enhancement and subsequent retrieval of fine resolution SM

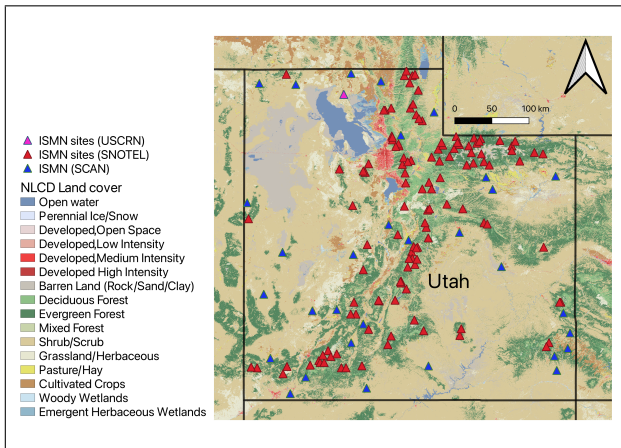


Fig. 2. ISMN sites with underlying land cover from MRLC National Land Cover Data (NLCD) for the year of 2021

2. METHODOLOGY

Our approach for enhancing the resolution of SM includes a physics-based and state-of-the-art deep-learning-based integrative framework. The cGAN [11] based framework is implemented for this downscaling purpose. This architecture adapts the image-to-image translation [12] to *disaggregate* 9 km TB to 1 km TB. This image translation contains U-NET as the *Generator* model in the cGAN architecture. The *Tau-omega* model is augmented with this U-NET to retrieve SM from TB. The intermediate output of the U-NET network is the 1 km TB, from which 1 km SM is retrieved using the physics-based model. U-NET is additionally constrained with

another DNN, which acts as the adversarial *Discriminator* model in the cGAN architecture. This DNN compares the predicted-retrieved 1 km SM with the ground truth (upscaled 1 km) SM and produces a classification score on whether the prediction is realistic. Using this classification the generator learns to mimic realistic predictions. The cGAN model learning is achieved by these two adversarial networks. During the learning process, for the discriminator, binary cross entropy loss would be optimized, whereas for the generator, both binary cross entropy and mean squared error losses would be optimized. This overall downscaling framework is shown in Figure 1. The subsections 2.1, 2.2 and 2.3 detail each of the Neural Network and physics based modules.

2.1. Generator model

The Generator model of the cGAN architecture consists of a CNN based modified U-NET network [9]. The U-NET consists of an encoder and a decoder block. Each encoder block consists of convolutional layers, followed by batch normalization layers passed through a non-linear Rectified Linear Unit (ReLU) activation function to get a real positive output. Each decoder block consists of transposed convolutional layers, followed by batch normalization layers, passing through a ReLU activation function. Additionally in the decoder block, a dropout layer is added to reduce overfitting during training. A modification of adding extra transposed convolutional layers to the U-NET architecture, to increase the spatial resolution, is made. This makes sure the output image size is 9 times the input image size, indicating resolution enhancement to 1 km. The input to the generator is coarse-resolution TB, while the output is fine-resolution TB.

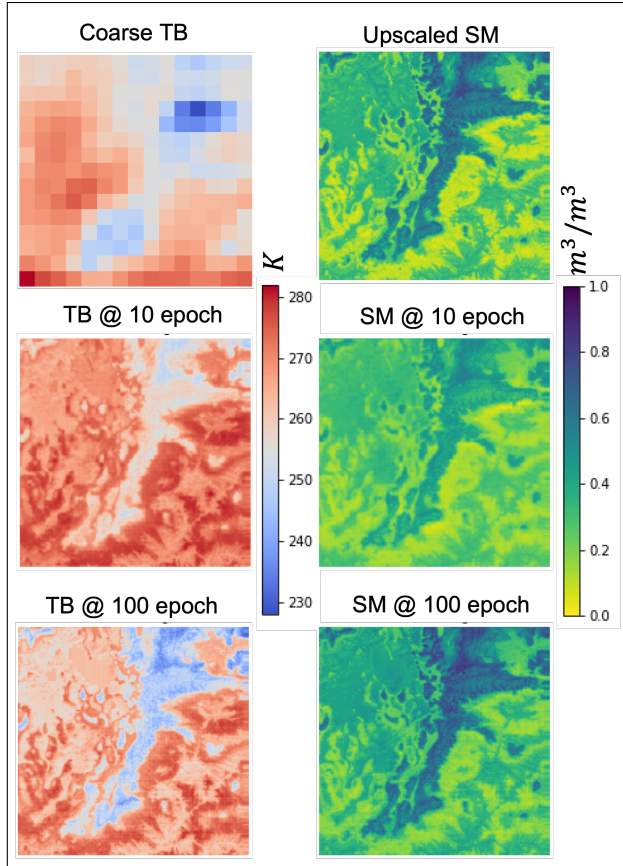


Fig. 3. Validation dataset TB and SM prediction at various training epochs

2.2. *Tau-omega* passive retrieval model

The predicted 1 km TB is then taken as input to the well-known *Tau-omega* model. Using the Single Channel Algorithm with Vertical polarization (SCA-V) retrieval method [13], 1 km SM is retrieved. The ancillary data (such as soil properties, vegetation properties) to the *Tau-omega* model are from the SMAP Enhanced L3 data product [8]. The input to the *Tau-omega* model is the fine-resolution TB from U-NET and the output is fine-resolution SM.

2.3. Discriminator model

When using Neural Networks for various applications, typical loss functions such as mean absolute difference, root mean squared errors, and image-based metrics such as peak signal-to-noise ratio, would be optimized while training. These loss functions do not allow sufficient flexibility in the prediction process to quantify the difference between predicted and observed images, especially in applications where we target 1 km per pixel TB estimation from 9 km data. This procedure requires a resolution scaling factor of 9 that the U-NET architecture has to learn from the training data. In order to ac-

commodate this large scaling, the U-NET (generator model) architecture was augmented with another DNN, serving as a Discriminator in the cGAN framework.

The Discriminator part of the cGAN acts as a classifier that distinguishes predicted-retrieved SM and true-upscaled SM. The discriminator model consists of convolutional layers and batch normalization layers, passed through a sigmoid activation function. This activation function ensures the output to be between 0 and 1, with 0 indicating that the given two input images are highly varied, and 1 meaning that both input images are the same. This model helps in *discriminating* the predicted vs true images.

Depending on the output classification, the generator model tweaks itself to reproduce images similar to that of up-scaled SM. This continuous learning by both models ensures that the predicted-retrieved SM is as close as possible to the ground truth (upscaled) SM.

3. STUDY AREA AND DATA

For the downscaling framework, the state of Utah, USA centered at 39.3210° N, 111.0937° W, was chosen to be the study area. Utah has relatively heterogeneous land cover types, comprising of shrublands, grasslands, and mixed forests, and has a dry, semi-arid and desert climate. This area contains the USCRN, SNOTEL, and SCAN networks all accessible from the International Soil Moisture Network (ISMN) [14]. This is illustrated in Figure 2.

The training data for the U-NET system to downscale SMAP L3 Enhanced 9 km TB to 1 km TB was generated using the ISMN *in-situ* SM measurements. The point SM measurements were first upscaled using the Random Forest model [15] using various ancillary input layers mentioned in [16]. Using the *Tau-omega* forward model [13], 1 km TB was further estimated from upscaled 1 km SM using the mosaicked SMAP L3 ancillary data (clay, albedo, surface temperature, roughness coefficient, vegetation optical depth) products. These coarse and fine resolution TB images, along with upscaled SM, are the inputs to the cGAN framework. The output of this network is fine resolution SM.

4. RESULTS AND DISCUSSIONS

We considered the satellite observations and *in-situ* measurements of year 2021 for the experiment. The entire dataset was split into 80% training, 10% validation, and 10% test datasets. The coarse resolution TB image is of size 55 × 55 at 9 km spatial resolution per pixel and the output fine resolution TB and subsequent SM are of size 495 × 495 at 1 km spatial resolution per pixel.

The predicted-retrieved SM image of a TB image from validation set at various training epochs is shown in Figure.3. We can see that cGAN model learned the SM spatial variability very well after 100 epochs.

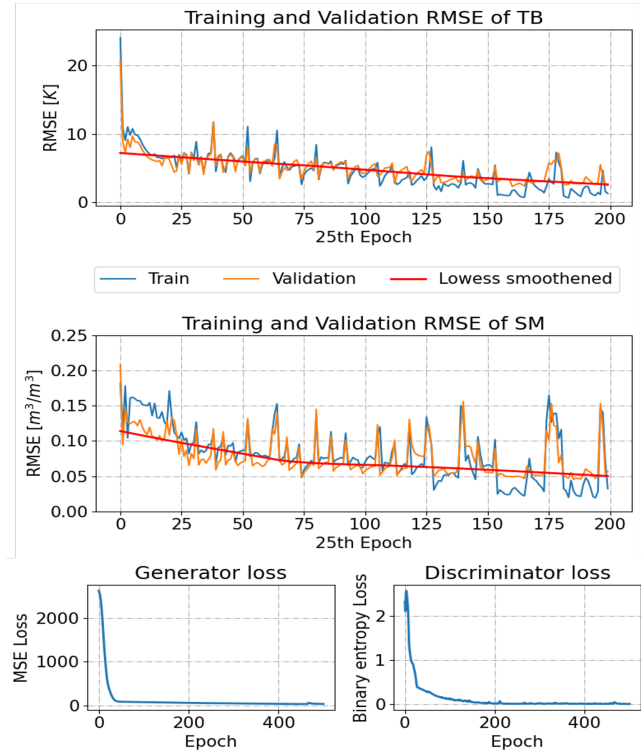


Fig. 4. RMSEs (top 2 plots), and model losses (bottom plots)

Data	Training	Validation	Testing
TB[K]	3.1255	3.3125	3.4191
SM[m³/m³]	0.0632	0.0690	0.0776

Table 1. Different datasets RMSE of predicted TB and SM

Figure 4 illustrates the Root Mean Squared Error (RMSE) of predicted fine resolution TB and SM against true images. The figure also shows the Locally Weighted Scatterplot Smoothing (LOWESS) plot, to depict the decreasing trend of RMSE as the training epochs increases. The bottom subplots illustrate the loss functions of two neural networks. We see a convergence from both the models. The RMSE of predictions of models with lowest validation error is tabulated in Table 1. Further error reduction would be possible with more training data.

5. CONCLUSION

Downscaling satellite images to high resolution images is a challenging task. In this work, we developed a cGAN framework coupled with a physics based model to downscale SMAP soil moisture and validated the results against up-scaled *in-situ* soil moisture. In future work, introducing radar backscatters to the brightness temperature and also ways to improve the ancillary datasets will be explored.

6. REFERENCES

- [1] S. Sabaghy, J. P. Walker, L. J. Renzullo, R. Akbar, S. Chan, J. Chaubell, N. Das, R. S. Dunbar, D. Entekhabi, A. Gevaert, T. J. Jackson, A. Loew, O. Merlin, M. Moghaddam, J. Peng, J. Peng, J. Piepmeier, C. Rüdiger, V. Stefan, X. Wu, N. Ye, and S. Yueh, "Comprehensive analysis of alternative downscaled soil moisture products," *Remote Sensing of Environment*, vol. 239, 3 2020.
- [2] J. Peng, A. Loew, O. Merlin, and N. E. Verhoest, "A review of spatial downscaling of satellite remotely sensed soil moisture," *Reviews of Geophysics*, vol. 55, pp. 341–366, 6 2017.
- [3] S. I. Seneviratne, T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, "Investigating soil moisture–climate interactions in a changing climate: A review," *Earth-Science Reviews*, vol. 99, no. 3–4, pp. 125–161, 2010.
- [4] N. N. Das, D. Entekhabi, R. S. Dunbar, M. J. Chaubell, A. Colliander, S. Yueh, T. Jagdhuber, F. Chen, W. Crow, P. E. O'Neill, *et al.*, "The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product," *Remote Sensing of Environment*, vol. 233, p. 111380, 2019.
- [5] H. Sun and Y. Cui, "Evaluating downscaling factors of microwave satellite soil moisture based on machine learning method," *Remote Sensing*, vol. 13, pp. 1–16, 1 2021.
- [6] W. Xu, Z. Zhang, Z. Long, and Q. Qin, "Downscaling SMAP Soil Moisture Products with Convolutional Neural Network," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 4051–4062, 2021.
- [7] Y. Fang, L. Xu, Y. Chen, W. Zhou, A. Wong, and D. A. Clausi, "A Bayesian Deep Image Prior Downscaling Approach for High-Resolution Soil Moisture Estimation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 4571–4582, 2022.
- [8] P. O'Neill, S. Chan, E. Njoku, T. Jackson, R. Bindlish, J. Chaubell, and A. Colliander, "SMAP enhanced L3 radiometer global and polar grid daily 9 km ease-grid soil moisture version 5," *NASA National Snow and Ice Data Center Distributed Active Archive Center (Ed.)*, Boulder, Colorado USA. DOI, vol. 10, 2021.
- [9] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," 2015.
- [10] T. J. Jackson and T. J. Schmugge, "Vegetation Effects on the of Soils Microwave Emission," *Remote Sensing Environment*, vol. 36, pp. 203–212, 1991.
- [11] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014.
- [12] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," 2016.
- [13] P. O'Neill, R. Bindlish, S. Chan, J. Chaubell, E. Njoku, and T. Jackson, "Soil Moisture Active Passive (SMAP) Algorithm Theoretical Basis Document Level 2 & 3 Soil Moisture (Passive) Data Products," *Jet Propulsion Laboratory*, 2020.
- [14] W. Dorigo, W. Wagner, R. Hohensinn, S. Hahn, C. Paulik, A. Xaver, A. Gruber, M. Drusch, S. Mecklenburg, P. van Oevelen, *et al.*, "The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements," *Hydrology and Earth System Sciences*, vol. 15, no. 5, pp. 1675–1698, 2011.
- [15] L. Breiman, "Random Forests," *Machine learning*, vol. 45, pp. 5–32, 2001.
- [16] D. Clewley, J. B. Whitcomb, R. Akbar, A. R. Silva, A. Berg, J. R. Adams, T. Caldwell, D. Entekhabi, and M. Moghaddam, "A method for upscaling in situ soil moisture measurements to satellite footprint scale using Random Forests," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 6, pp. 2663–2673, 2017.