

## Physics-Guided Machine Learning Based Forward-Modeling of Radar Observables a Case Study on Sentinel-1 Observations of Corn-Fields

Nikaein, Tina; Lopez-Dekker, Paco

**DOI**

[10.1109/JSTARS.2025.3543238](https://doi.org/10.1109/JSTARS.2025.3543238)

**Publication date**

2025

**Document Version**

Final published version

**Published in**

IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

**Citation (APA)**

Nikaein, T., & Lopez-Dekker, P. (2025). Physics-Guided Machine Learning Based Forward-Modeling of Radar Observables: a Case Study on Sentinel-1 Observations of Corn-Fields. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18, 6492-6502.  
<https://doi.org/10.1109/JSTARS.2025.3543238>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Physics-Guided Machine Learning-Based Forward-Modeling of Radar Observables: A Case Study on Sentinel-1 Observations of Corn-Fields

Tina Nikaein<sup>✉</sup>, Member, IEEE, and Paco Lopez-Dekker<sup>✉</sup>, Senior Member, IEEE

**Abstract**—Artificial neural networks have the potential to model the interaction of radar signals with vegetation but often do not follow the physical rules. This article aims to develop a new physics-guided machine learning approach that combines neural networks and physics-based models to leverage their complementary strengths and improve the modeling of physical processes. We propose a data-driven framework to model synthetic aperture radar observables by incorporating physical knowledge in two ways: through the network architecture and the loss function. A key aspect of our approach is its ability to integrate knowledge encoded in physics-based models. The results show that by using scientific knowledge to guide the construction and learning of the neural network, we can provide a framework with better generalizability and stability.

**Index Terms**—Constraint, forward model, neural network, physics-guided, radar backscatter, synthetic aperture radar (SAR).

## I. INTRODUCTION

**F**ORWARD models, or observation operators, are important for the analysis and interpretation of remote sensing data, for the conceptualization and development of observational concepts, and for the assimilation of measurements in numerical models. As in many other fields of study, traditional approaches use either simple empirical models, simplified physical models, or a mixture of both (e.g., physical models with empirically tuned parameters). While robust and easy to work with, these types of models often fail to account for many of the phenomena present in the full physical system.

To address this shortcoming, remote sensing scientists are increasingly adopting machine learning (ML) algorithms. These algorithms can learn complex relations and patterns that are not well captured by theoretical models due to the complexity of the underlying physics [1], [2]. On the other hand, supervised ML algorithms, particularly deep learning networks, often require vast amounts of training data and can sometimes yield results that, while statistically accurate, may not always align with physical laws [3]. Another issue with purely data-driven models is that they may work well for the region of input vector-values

covered by the training data, but often fail to generalize for input values outside this region [4].

An emerging trend in physical sciences is to use the robust theoretical foundations of physics to guide and constrain ML models, leading to predictions that are not only more accurate, but also physically plausible. In this work, we follow this approach to model satellite-based radar observables over crop fields. While we apply the methodology to the particular case of Sentinel-1 [5] Normalized Radar Cross Section (NRCS) over corn fields, the approach should apply to other crops and other observables.

Recently, there has been an increasing interest in the integration of physics with machine learning, as discussed in detail in [6] and [7]. Previous studies demonstrated that physics-informed machine learning can improve the accuracy and generalizability of the model in different applications. For example, in [8] an improvement in the prediction of the chemical reflectance signature was studied using a physics-guided neural network (PGNN). In another study to predict lake temperature [9], PGNN was used in two different approaches: 1) the simulated output of the physics-based model was fed into the neural network as additional inputs; and 2) including physics knowledge into the loss function. Their results showed better accuracy and lower physical inconsistency. Jia et al. [10] pretrained a model using simulated data from a generic physics-based model to improve prediction accuracy with limited observed data. An effective method for guiding the initialization process to aid in model training and avoiding local minima is to employ transfer learning, an ML approach. With transfer learning, a model can be first pretrained using simulated data from a physics-based model and subsequently fine-tuned with a limited amount of training data to adapt to the specific task at hand. Their results show that using physical model data for pretraining, even with imperfect parameters, can reduce the training data requirements. They incorporate the knowledge encoded in the physical model with a recurrent neural network model to leverage their complementary strengths to predict lake water temperature. Zhong et al. [11] developed a physics-informed deep learning model to simulate runoff changes in alpine catchments under climate change, outperforming traditional models. Their model combines deep learning techniques with the physics of hydrological processes, providing more credible projections. Previous studies showed the efficiency of including physical knowledge into the architecture of the model, for example in [12]

Received 7 November 2024; revised 12 January 2025; accepted 14 February 2025. Date of publication 18 February 2025; date of current version 6 March 2025. (Corresponding author: Tina Nikaein.)

The authors are with the Department of Geoscience and Remote Sensing, Delft University of Technology, 2628 CN Delft, The Netherlands (e-mail: tnikaein@tudelft.nl).

Digital Object Identifier 10.1109/JSTARS.2025.3543238

and [13], where prior system knowledge was incorporated in the architecture, for lake temperature modeling and dynamic system modeling, respectively. In all state-of-the-art applications, the integration of ML with physical knowledge has led to significant improvements in adaptability. To the best of our knowledge, the application of these advancements to modeling microwave signals from vegetation remains unexplored. This gap presents a unique opportunity for research, where methodologies developed in other contexts could be adapted to enhance the accuracy and efficiency of remote sensing in vegetative environments.

The advantages of direct assimilation of microwave satellite observation, which circumvents the need for retrievals, were studied in [14] and [15]. Forward models are required to map biogeophysical parameters to satellite observations. In the context of this article, our aim is to model the dependence of NRCS,  $\sigma_0$ , on crop and soil-related parameters, in order to use this model for direct assimilation. A commonly used radiative transfer model is the so-called water cloud model (WCM) [16], which often serves as a forward operator. There are more complex models to simulate radar backscatter, like the Tor Vergata model [17] and the Michigan microwave canopy scattering model [18]. While these models offer detailed simulations, their use is often limited due to the difficulty in parameterizing them accurately. These models require a large number of input parameters, many of which are challenging to measure or estimate with high precision in real-world conditions. More recently, the advantage of using machine learning as an observation operator has been studied by [19], [20], [21], and [22]. The challenge in accurately predicting NRCS lies in the complex interplay of numerous factors influencing the returned signal, including surface roughness, moisture content, vegetation cover, and geometric properties of the observed scene. The study by Nikaen et al. [22], highlights the difficulties in using data-driven models alone to simulate synthetic aperture radar (SAR) observables, such as backscatter, during anomalous conditions, e.g., drought years. This challenge arises when the model encounters scenarios for which it has not been trained, such as vegetation parameters under dry conditions, resulting in predictions that are not representative of the actual conditions. This research builds on these insights and seeks to address these limitations by demonstrating how the integration of physical knowledge into ML models can improve their performance as an observation operator. Specifically, the focus is on the development of physics-guided machine learning frameworks that incorporate domain knowledge into both the architecture and learning process of ML models. By doing so, this approach not only ensures that the predictions remain consistent with fundamental physical principles, but also enhances the ability of the model to generalize to previously unseen conditions, such as drought or other environmental anomalies. We used machine learning as an observation operator to map biogeophysical parameters from crop growth models, such as the Decision Support System for Agrotechnology Transfer (DSSAT), to SAR observables. This approach provides a more robust framework for simulating SAR observables.

The main aim of the article is to investigate how incorporating physical constraints into ML models can enhance their robustness, generalizability, and interpretability. Rather than focusing

on developing the best-performing ML model, our objective is to understand how physical principles can be embedded into the modeling process. To achieve this, we explore two complementary approaches. The first approach incorporates physical knowledge directly into the learning process by adding a custom constraint to the loss function. A positive gradient constraint enforces consistency with the expected relationship between radar backscatter and soil moisture. The second approach mirrors the structure of the WCM, decomposing radar backscatter into physically meaningful components (soil and vegetation) and constraining their contributions to the total signal. This implementation enables explicit modeling of attenuation effects, providing deeper insights into the physical processes governing SAR observables.

The contributions of this work are multifold and are as follows.

- 1) We demonstrate that incorporating physical constraints improves model robustness and transferability across different environmental conditions, such as year-to-year variability.
- 2) We highlight the potential of gradient-based constraints, as an effective way to embed physical principles into ML models for modeling radar observables.
- 3) We show that mirroring the WCM structure in the neural network architecture allows for intermediate outputs that are physically interpretable, enabling the analysis of specific contributions from soil and vegetation to backscatter.
- 4) By using synthetic data derived from the WCM, we validate the accuracy and behavior of the proposed models under controlled conditions, providing a benchmark for real-world applications.

This article provides a practical framework for integrating physical principles into ML models, with the dual goal of improving performance and gaining deeper insights into the underlying processes. The findings contribute to advancing the state of physics-informed machine learning and its applications in remote sensing.

## II. STUDY AREA AND DATA

Building on the work presented in [22], we select maize fields in the province of Noord-Brabant, The Netherlands for our study. The Crop Environment Resource Synthesis (CERES)-Maize model, which is among the various crop models included in the DSSAT [23], [24], were used to simulate crop growth for each field. This model uses input data on soil characteristics, climatic conditions, crop genetics, and management practices to simulate daily growth stages, biomass development, and crop yield. Our research focuses on key biophysical parameters of maize, such as the Leaf Area Index (LAI), Above-Ground Biomass (AGB), surface soil moisture ( $SM_S$ ), and root zone soil moisture ( $SM_R$ ), to simulate SAR observables. More details, including details about the study area and the crop growth modeling steps, can be found in [22].

Following [22], we utilize Sentinel-1 C-band data acquired in the Interferometric Wide Swath mode with a six-day repeat cycle (relative orbit 37). NRCS values in both VV and VH

polarizations were retrieved from the Agricultural SandboxNL database [25]. In order to validate the rationale of our proposed approach, we generate synthetic data based on the principles of WCM. This synthetic data allows us to test and validate the accuracy, robustness, and ability of the model to generalize across different scenarios in an idealized case.

### III. METHODOLOGY

In this section, we describe the two main approaches to incorporate physical knowledge into neural networks: 1) physics-based loss function; and 2) physics-guided network topology.

#### A. Physics-Based Loss Function

One way to incorporate physical knowledge into ML algorithms is by enforcing constraints on the outputs of these algorithms. This can be accomplished through the formulation of a custom loss function, as shown in [9]. Consider a learning system characterized by a function  $f$ , which operates on a set of input parameters  $\mathbf{X}$  that possess a physical relationship with the target variable  $Y$ . In this context, we can express the relationship as  $\hat{Y} = f_{\text{NN}}(\mathbf{X})$ , where  $f_{\text{NN}}$  denotes the neural network function approximating the mapping from  $\mathbf{X}$  to  $\hat{Y}$  across our training samples. In the conventional training paradigm, the goal is to minimize the discrepancy between the predicted values ( $\hat{Y}$ ) and the observed values ( $Y$ ). However, while this standard approach is effective in reducing predictive error, it may not ensure that predictions are in accordance with the underlying physical principles. To address this gap, the custom loss function comes into play, integrating physical constraints directly into the learning process. By doing so, the loss function not only penalizes deviations from observed data points, but also incorporates penalties for violations of known physical laws. This dual-purpose loss function ensures that the learning process is not merely data-driven but is also guided by the underlying physical principles. The modified learning objective, incorporating this physical constraint, is defined as

$$\arg \min_f \left( L_{\text{data}}(Y, \hat{Y}) + \lambda L_{\text{phys}}(\hat{Y}) \right)$$

with

$$L_{\text{data}}(Y, \hat{Y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

where  $f$  is the model,  $L_{\text{data}}$  is the data term of the loss function,  $L_{\text{phys}}$  is the physical constraint term, and  $\lambda$  is a hyperparameter that balances the contributions of  $L_{\text{data}}$  and  $L_{\text{phys}}$  to the overall loss function. For the data term of the loss function, we use the mean square error (MSE) function. To operationalize this physical constraint within our machine learning model, we developed a custom loss function with an additional term specifically designed to ensure the partial derivative of backscatter with respect to surface soil moisture remains positive [26], as

$$L_{\text{phys}}(\hat{Y}) = \frac{1}{n} \sum_{i=1}^n \text{ReLU} \left( -\frac{\partial \hat{y}_i}{\partial \text{SM}_s} \right) \quad (2)$$

where index  $i$  iterates over the training samples, and  $\text{ReLU}(\cdot)$  is a Rectified Linear Unit function applied within the  $L_{\text{phys}}$  term to enforce the positive partial derivative constraint by penalizing negative values of the predicted partial derivative.

This adjustment is critical for maintaining the physical integrity of the predictions of the model, ensuring they are consistent with the known behavior of microwave radar signals interacting with varying levels of soil moisture.

#### B. Physics Guided Network Topology

In this case, we are constraining the internal architecture of ML models with physical insights to enhance their interpretability. This approach involves integrating physical principles directly into the structure of neural network, as shown in Fig. 1. By doing so, we give an implicit physical meaning to some intermediate outputs, which also means that we can apply physical constraints to them. In this framework, we try to limit the existing freedom of a standard neural network to simulate NRCS. For this approach, we tried two steps: 1) bound the freedom of the model in the architecture; and 2) incorporate the physical knowledge through the loss function. The network topology tested follows the architecture of the WCM [16], a widely used model for backscatter. Here, the total NRCS during the growth period (from planting to harvest) is decomposed in a vegetation component, and underlying soil term, and a term representing their interaction. The general form of the WCM equations is represented in (3) to (6), where the WCM neglects the interactions between the ground and vegetation, implicitly assuming that it is small compared to the other terms

$$\sigma_{\text{total}}^0 = \sigma_{\text{veg}}^0 + \sigma_{\text{soil}}^0 T^2, \quad (3)$$

$$\sigma_{\text{veg}}^0 = A V_1 \cos \theta (1 - T^2) \quad (4)$$

$$T^2 = \exp \left( \frac{-2BV_2}{\cos \theta} \right) \quad (5)$$

$$\sigma_{\text{soil}}^0 = C + D \cdot \text{SM}_s. \quad (6)$$

Here,  $\theta$  is the incidence angle,  $\sigma_{\text{total}}^0$  is the total backscattering coefficient,  $\sigma_{\text{veg}}^0$  is the backscatter contribution from the vegetation,  $\sigma_{\text{soil}}^0$  is the backscatter contribution from the soil, and  $T^2$  is the two-way transmissivity of the vegetation layer. There are more sophisticated ways to estimate  $\sigma_{\text{soil}}^0$  (e.g., [27]) but generally, it is a reasonable assumption that there is a linear relationship between backscattering coefficient and soil moisture over bare soil.  $\sigma_{\text{soil}}^0$  influenced by soil moisture, surface roughness, and the incidence angle of the radar signal. The attenuation term depends on the density and water content of the vegetation. Scattering from the vegetation depends on vegetation water content, structure, and orientation of leaves and stems [28]. The WCM contains four coefficients, A, B, C and D related to vegetation scattering, vegetation attenuation, surface roughness, and soil moisture respectively. All of these coefficients are polarization- and frequency-dependent. A and B depend on vegetation type, while C and D are related to soil texture. Several quantities can be used to describe the vegetation by setting  $V_1$  and/or  $V_2$  to quantities such as vegetation water content (VWC), vegetation



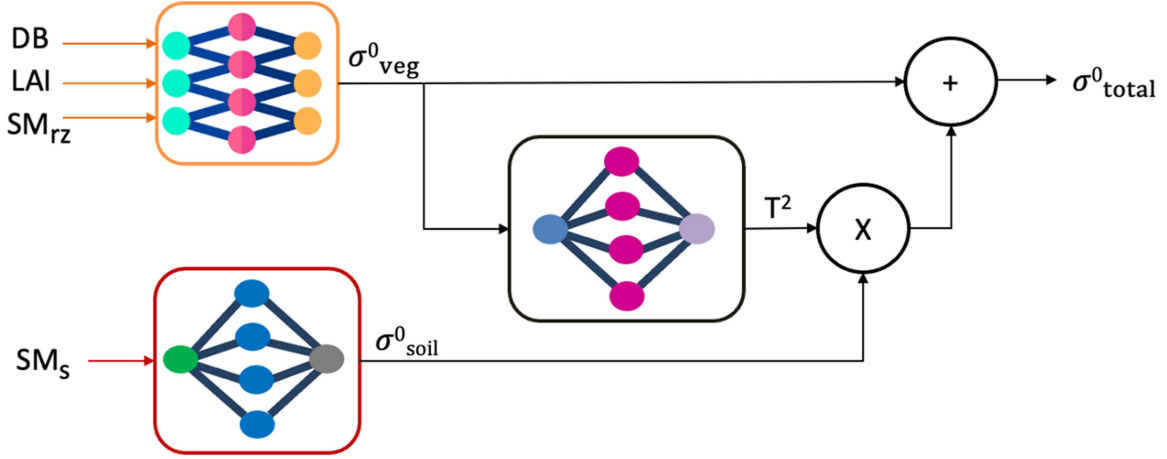


Fig. 1. A depiction of the WCM-inspired NN, where prior knowledge from physics is embedded in a structured format.

optical depth, AGB, or LAI. As LAI is readily available from the DSSAT model outputs, we follow [29] and [30] by assuming  $V_1 = V_2 = \text{LAI}$ .

We defined a multiple input network that separates the inputs for the vegetation component from the inputs for the soil component. The vegetation components inputs are LAI, AGB, and  $\text{SM}_R$  (as an indication of the availability of root zone moisture to replace transpired water) and  $\text{SM}_S$  is the input for the soil component. As illustrated in Fig. 1, the network produces internal outputs denoted as  $\sigma_{\text{veg}}^0$  and  $\sigma_{\text{soil}}^0$ . We anticipate that any increase in  $\sigma_{\text{veg}}^0$  will result in a diminished direct impact from  $\sigma_{\text{soil}}^0$ . This modulation effect is represented by the term  $T^2$  within the framework and multiplied by NRCS from the soil term. Note that in the attenuation term of the constrained model architecture, we prescribe a behavior inspired by the reverse exponential nature of attenuation and vegetation effects as it is shown in (7). This term has a trainable coefficient  $\alpha$ , which is defined by the training procedure.

$$T^2 = \exp(-\alpha \sigma_{\text{veg}}^0). \quad (7)$$

The proposed framework incorporates physics-based prior knowledge into the structure of the neural network and imposes constraints on the internal states of the model and the output values. In this architecture, the loss function constraint is applied to  $\sigma_{\text{soil}}^0$  rather than  $\sigma_{\text{total}}^0$ . This choice is based on the observation that the partial derivative with respect to  $\text{SM}_S$  for  $\sigma_{\text{soil}}^0$  is consistently positive. The sensitivity of  $\sigma_{\text{total}}^0$  to  $\text{SM}_S$  can approach zero when the LAI is high, so in the presence of noise, this can lead to excessive penalties for slightly negative values.

The rest of this article is devoted to analyzing the outcomes, testing, and comparing the four combinations discussed: Standard NN and WCM-inspired NN, each with either a regular or modified loss function. As mentioned in Section II, synthetic data were generated using the WCM formula, which allows the extraction of intermediate layer output. This data enables us to validate the performance of the WCM-inspired NN model. In this study, we formulate three key hypotheses that we aim to test and validate through our results as follows.

- 1) When training and testing on consistent data sets, adding any constraints would lead to a higher MSE, given that this is our default (unconstrained) loss function.
- 2) Both our positive partial derivative and topology constraint should enhance the robustness of the model and, consequently, its transferability.
- 3) We anticipate that our WCM-inspired neural network will learn to produce meaningful intermediate results.

### C. NN Implementation

We implemented the neural network models using the Keras library [31]. The dataset was partitioned field-wise into training and testing subsets, following a 70:30 ratio to ensure independence between the two datasets. We used the Adaptive Moment Estimation (Adam) optimization algorithm to minimize the loss function. To mitigate the risk of overfitting, we incorporated an early stopping mechanism, adjusting the patience parameter to 50. The input features were rescaled between 0 and 1 to prevent saturation at the tails of the activation functions; the same transformation was applied to the test data. Our fully connected network architecture consists of 3 hidden layers and neuron numbers of 32, 16, 4 with ReLU activation functions in each hidden layer. In the constrained model architecture, ReLU was used in the hidden layers, while the sigmoid activation function was used in each output layer to produce outputs between 0 and 1. The weights of the neural network were randomly initialized and each experiment was run 50 times where the standard deviation of accuracy was around 2%.

We present results and compare outcomes in the following scenarios: standard NN, standard NN with constraints, WCM-inspired NN, and WCM-inspired with constraints.

## IV. EXPERIMENT WORKFLOW

In this section, we present the results from each approach for two scenarios: 1) when the model is tested on data from the same year, incorporating the environmental conditions present in the training phase; and 2) when the model is tested on data from

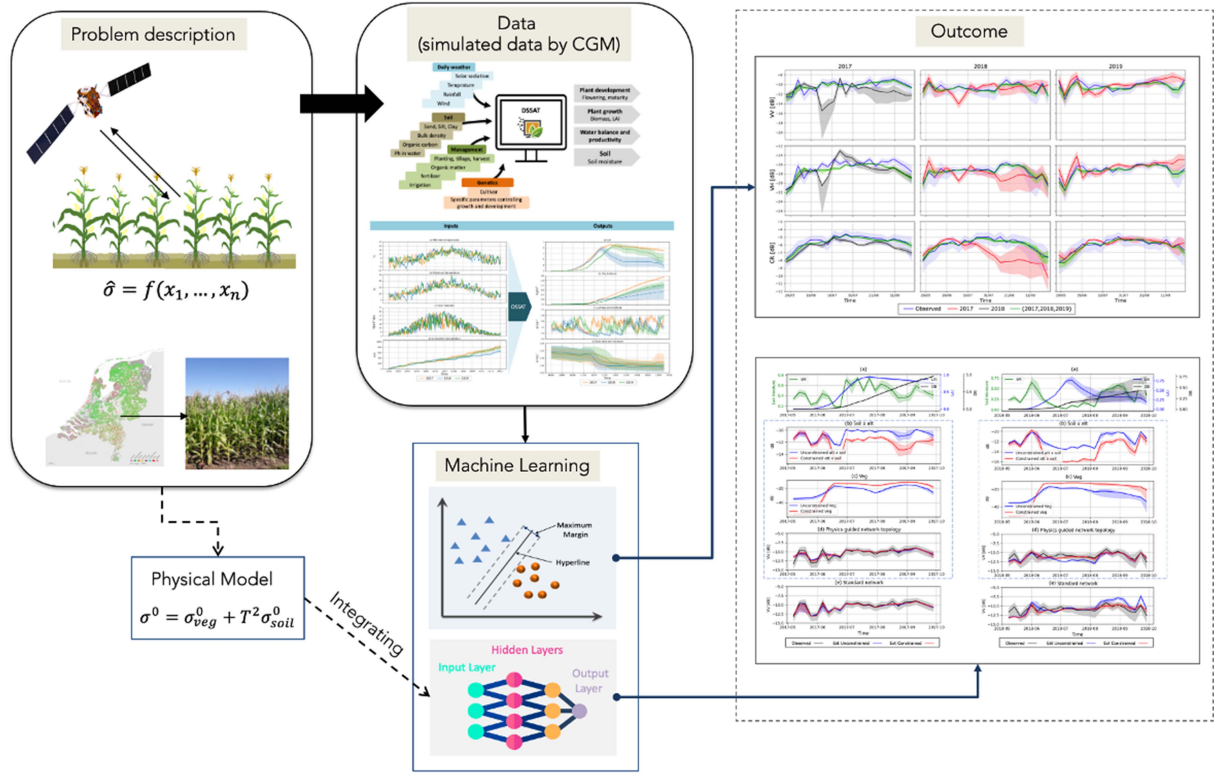


Fig. 2. General workflow of our proposed framework. The simulated data from the crop growth model was fed to the model. The details of data simulation and modeling the NRCS using machine learning are discussed in [22]. (The map of The Netherlands is sourced from [25]).

a different year, which includes different environmental conditions not represented in the training data. The details of these environmental differences are detailed in [22]. Fig. 2 illustrates the general workflow of our proposed framework, demonstrating the simulation of NRCS using ML for corn parcels. The dashed arrows in this figure highlight the added value of this method to the data-driven study, enhancing it with physical principles not addressed in our previous work [22].

As mentioned in the previous section, these models aim to improve transferability. Specifically, the proposed method is tested in two distinct situations: 1) training and testing on the same year; and 2) testing on a different year with varying meteorological conditions. The lack of separate signals for soil and vegetation in satellite data limits the ability to validate the model effectively. To address this limitation, synthetic data can be used to understand better whether the model behaves as expected. We expect that the model performs optimally under these controlled conditions. We generate synthetic data by using soil and vegetation states modeled by DSSAT for both years as inputs to the WCM. The values for WCM parameters were optimized by the range provided in [21].

## V. RESULTS

### A. Synthetic Data

Fig. 3 shows the results obtained by training the physics-guided model on synthetic data for 2017, testing on data for the same year Fig. 3(i) and on data for 2018 Fig. 3(ii). This

figure illustrates how constraints on the network topology by physical properties such as soil moisture and vegetation characteristics can improve remote sensing models' accuracy. The figure provides insight into the contributions to total backscatter. Fig. 3(ii)(a) illustrates a rapid decrease in LAI during July, along with lower maximum AGB and greater variability across the parcels (indicated by the shaded area). In Fig. 3(ii)(b),  $T^2 \sigma_{soil}^0$  is overestimated in late-June/early-July; however, this has limited impact on  $\sigma_{total}^0$  since  $\sigma_{veg}^0$  dominates during this period. Similarly, while  $\sigma_{veg}^0$  is overestimated in June, it has a negligible effect on  $\sigma_{total}^0$  because  $\sigma_{veg}^0$  is much smaller than  $\sigma_{soil}^0$  at that time. The graph highlights that the largest deviation between the estimated and synthetic truth  $\sigma_{total}^0$  occurs from mid-July onwards, coinciding with an anomaly in the LAI, which leads to a poor estimate of  $\sigma_{veg}^0$ . However, this difference remains minimal, as shown in Fig. 3(ii)(d). The inclusion of a constraint in the loss function helps to reduce this deviation, bringing the estimate slightly closer to the synthetic truth, and improving the accuracy of the model despite the anomaly. In subplot (c), while the difference between synthetic (WCM) and estimated backscatter looks large prior to June 2017, note that in linear units, these initial values are all close to zero, so the absolute difference is very small.

The physics-guided network topology [Fig. 3(d)] with  $R^2 \approx 0.99$  and an MSE close to zero, effectively captures the interactions between the soil and attenuation [Fig. 3(b)] and the vegetation [Fig. 3(c)]. This integration is consistent with the principles of the WCM. The variability and trends observed

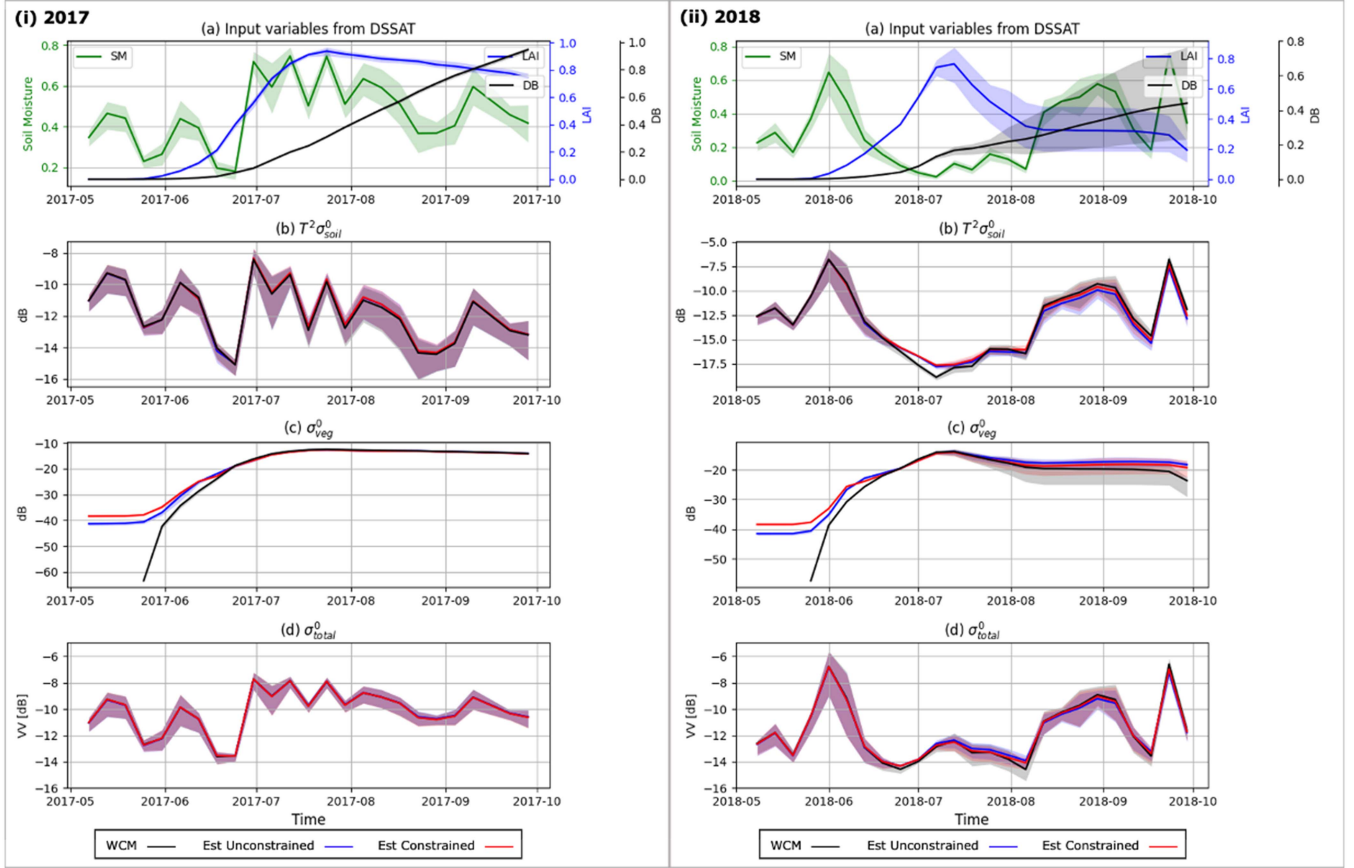


Fig. 3. Time series of synthetic and estimated NRCS values are presented, based on training using synthetic data generated from 2017 vegetation and soil parameters. The models were tested on (i) synthetic data from the same year (2017) and (ii) synthetic data generated from 2018 vegetation and soil parameters. (b) to (d) illustrate the following: Blue lines represent unconstrained estimations, red ones indicate constrained estimations, and the black lines represent the estimated backscatter from WCM.

in Fig. 3(d) are direct results of the dynamic changes in soil moisture and vegetation properties captured in Fig. 3(b) and (c). Maybe the most salient observation is that our WCM-inspired model produces accurate values of  $\sigma_{\text{soil}}^0$  and  $\sigma_{\text{veg}}^0$  without being trained with corresponding data. This happens because the WCM-inspired topology cannot *find* another way to minimize the MSE in a situation where the NN topology is a perfect match to the actual model. It would be risky to conclude that the same behavior will automatically happen with real data given the simplifying assumptions of the WCM.

### B. Sentinel-1 Data

After assessing the performance of the proposed method on synthetic data, we now proceed to evaluate its performance using satellite data. Fig. 4 provides a comparison of the NRCS in VV polarization between the two approaches that incorporate physical knowledge into neural networks, with each model trained and tested on data from 2017. Fig. 4(a) illustrates the simulated biogeophysical parameters of maize fields in 2017, which are important to understand the subfigures (b), (c), and (d). Fig. 4(b)–(d) show the  $T^2\sigma_{\text{soil}}^0$ ,  $\sigma_{\text{veg}}^0$ , and  $\sigma_{\text{tot}}^0$  generated by our trained WCM-inspired neural network. Fig. 4(e)

shows the total NRCS for the *regular* fully connected dense network.

Fig. 4(b) shows the contributions of NRCS from the interaction between soil and attenuation. The blue line corresponds to the model trained without physical constraints, while the red line corresponds to the model trained with the modified loss function. Both models follow a similar trend, but the constrained model shows more variability, particularly from early June onward. This suggests that the constrained model is more responsive to fluctuations in soil moisture. Fig. 4(c) shows the NRCS contributions from vegetation. Both models show an increasing trend from June, are aligned with the growing season, and follow the same behavior as LAI. Fig. 4(d) shows the overall NRCS simulated by the physics-guided network topology. The variability seen in subplot (b) due to soil and attenuation interactions directly impacts the overall NRCS in subplot (d). Peaks and troughs in soil moisture (subplots (a) and (b)) correspond to similar variations in the total NRCS, highlighting the sensitivity of the model to soil moisture dynamics. The increasing trend observed in subplot (d) from mid- to late-2017 aligns with the progression of the growth season, as indicated by increasing LAI and AGB in subplot (a). This demonstrates that the physics-guided network effectively

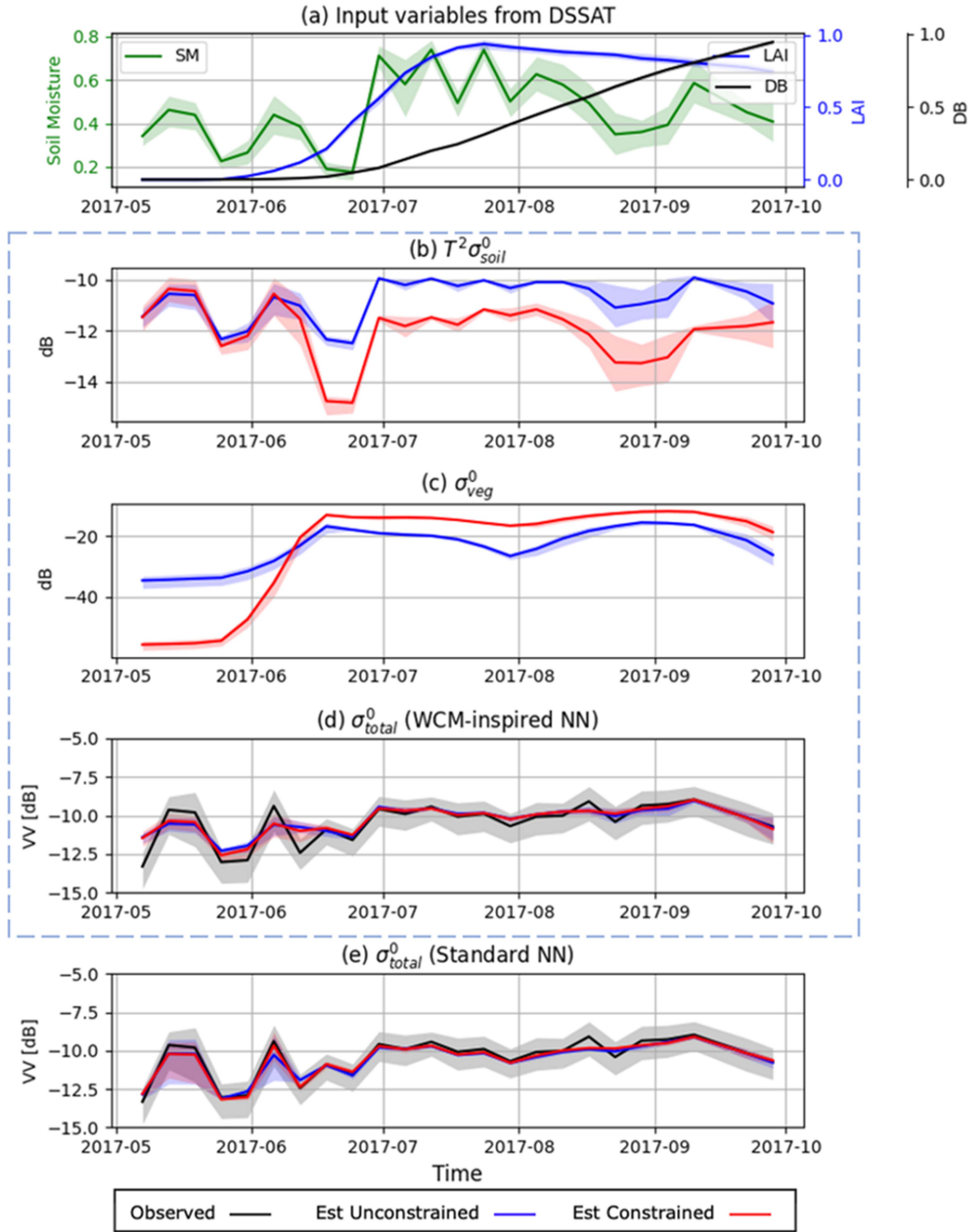


Fig. 4. Time-series of observed and estimated NRCS for training and test data on the same year data (2017). Solid lines indicate the mean value of the feature over maize parcels in the test set, and the bounded area shows the 20th-80th percentiles. (a) Vegetation and soil parameters during the growing season in daily steps in 2017. (b) to (e) illustrate the following: blue lines represent unconstrained estimations, red ones indicate constrained estimations, in (d) and (e), the black lines represent the observed backscatter. The dashed rectangle around subplots (b), (c), and (d) highlights that these are related to the physics-guided network topology method.

integrates the seasonal growth patterns of vegetation and changes in soil conditions. Fig. 4(e) illustrates the NRCS values obtained using a standard neural network. It compares the observed NRCS (black line) with estimates from the standard neural network without constraints (blue line) and with constraints (red line). The constrained model shows a closer alignment with the observed NRCS data. The constraint enforces a positive gradient with respect to soil moisture, improving the responsiveness of the model to changes in soil conditions. Interestingly, the total

NRCS from the WCM-inspired NN does not improve much with the additional constraint. However, the primary effect of incorporating the constraint into the loss function is to reduce the  $T^2 \sigma_{soil}^0$  term and increase  $\sigma_{veg}^0$  as soon as LAI begins to rise. Notably, throughout most of the growing season, the estimates of  $\sigma_{total}^0$  from both the WCM-inspired NN and the standard NN are quite similar. The notable exception occurs during the bare soil period, where the standard NN appears to perform slightly better.



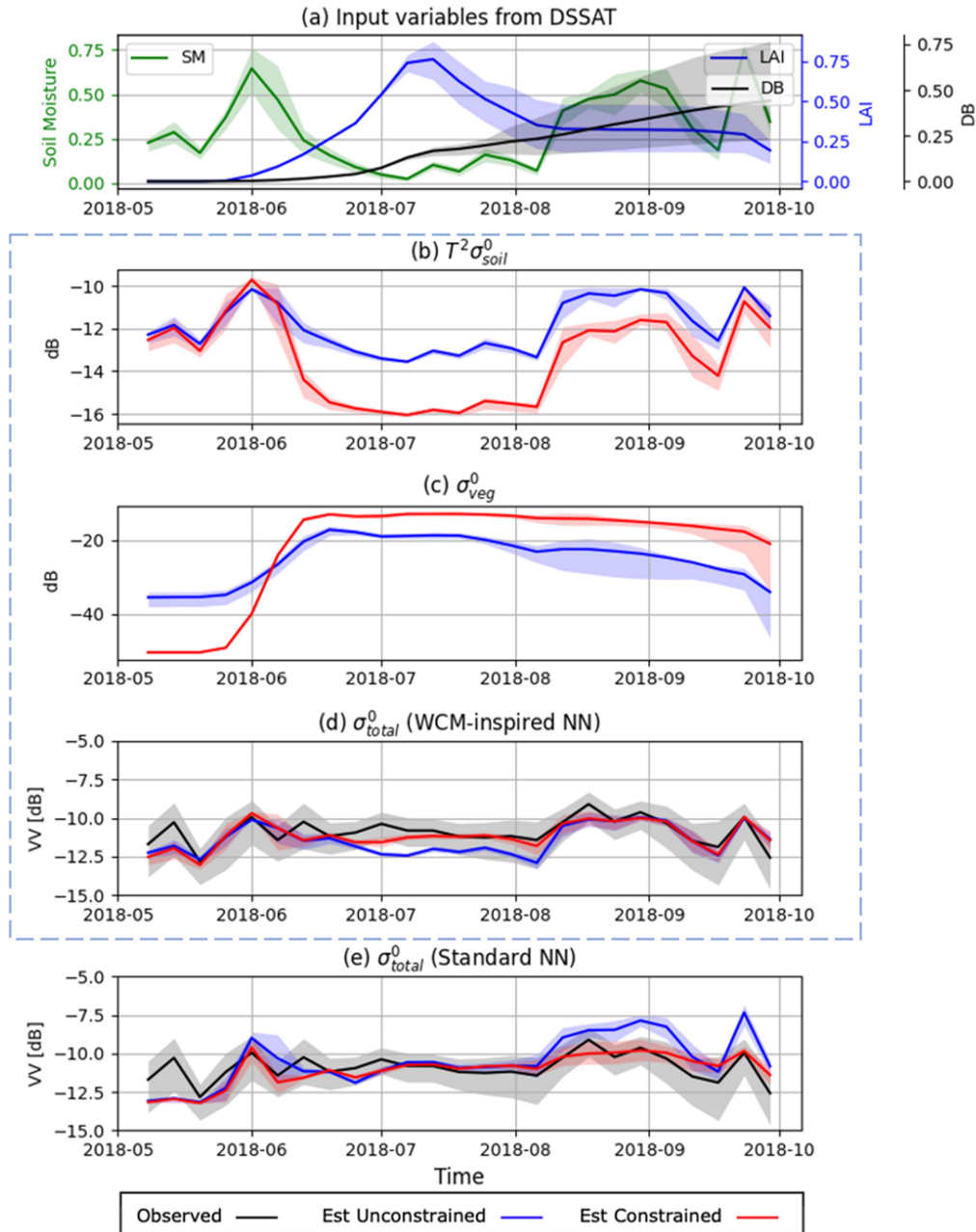


Fig. 5. Time-series of observed and estimated NRCS for training on 2017 data and testing on 2018 data. (a) Vegetation and soil parameters during the growing season in daily steps in 2018. (b) to (e) illustrate the following: blue lines represent unconstrained estimations, red ones indicate constrained estimations, in (d) and (e), the black lines represent the observed backscatter. The dashed rectangle around subplots (b), (c), and (d) highlights that these are related to the physics-guided network topology method.

Now we turn our attention to how the different models, trained on 2017 data, behave and perform on test data corresponding to 2018. As discussed before, the 2018 period includes combinations of input values not seen in 2017. This cross-year analysis helps to evaluate the transferability and robustness of the model under different environmental conditions.

Fig. 5 shows the model inputs and outputs in this case. Fig. 5(a) presents the DSSAT generated  $SM_s$ , LAI, and AGB over time in 2018. A comparison with Fig. 4(a) immediately reveals the severe drought during the summer months, as noted in [32], and its impact on the LAI and AGB. Panel (b) shows

that the constrained WCM-inspired model predicts significantly lower values for the  $T^2\sigma_{soil}^0$  term during the drought period. Subplot (c) shows that, as in the 2017 case, the constrained model produces higher contributions of the vegetation to the total NRCS and, more importantly, a sharper contrast between the bare-soil period and the growth period. The estimated NRCS without constraint in Fig. 5(d) follows the observed data but shows some deviations, especially between mid-June and August. The constrained model aligns well with the observed data, demonstrating improved accuracy in capturing NRCS variations due to changes in soil moisture and vegetation. The NRCS

TABLE I  
COMPARISON OF MODEL'S PERFORMANCE WITH AND WITHOUT CONSTRAINTS FOR TESTING ON DATA FROM 2017 AND 2018

Year	Model	Constrained	MAE	MSE	R2	Pearson	Spearman
2017	Standard	No	1.06	1.83	0.45	0.67	0.62
		Yes	1.03	1.76	0.47	0.69	0.63
	Topology	No	1.18	2.28	0.32	0.60	0.54
		Yes	1.14	2.16	0.35	0.62	0.58
2018	Standard	No	1.64	4.43	-0.35	0.34	0.33
		Yes	1.39	3.32	-0.01	0.30	0.31
	Topology	No	1.41	3.18	0.03	0.34	0.33
		Yes	1.33	2.94	0.10	0.38	0.36

values obtained using the standard neural network are shown in Fig. 5(e). In the standard NN, the estimate is poor during the LAI anomaly, and the constraint brings the estimate closer to the observed  $\sigma_{\text{total}}^0$ . When both NNs include a constraint in the loss function, the overall performance is better for the WCM-inspired NN.

Table I provides the performance metrics for different models for both years. As it was shown and discussed in [22], if the trained model is applied to sets of inputs corresponding to conditions (e.g., severe drought) for which it has not been trained we expect decreased model performance. The fully connected dense neural network demonstrates higher accuracy in predicting NRCS values when trained and tested on data from the same year. Although constraints improve the performance of the standard network, the improvement is less pronounced than in the physics-guided network case. When trained on 2017 data and tested on 2018 data, the WCM-inspired network predicts NRCS values with a higher accuracy. The correlation coefficients for 2018 are relatively low. However, it is crucial to emphasize that these metrics are calculated for the entire period under consideration. Notably, a significant improvement in these metrics is observed during the anomaly in late summer 2018. Constraints further improve both models, but the physics-guided network remains superior in terms of accuracy and robustness.

The results indicate that our physics-constrained method improves the forward modeling of SAR observables, demonstrating the potential to combine machine learning with scientific knowledge for advanced remote sensing applications.

## VI. CONCLUSION

In this article, we introduced a physics-guided neural network to model SAR observables over vegetation. Unlike traditional black-box neural networks, our approach integrates physical principles directly into the network architecture and the loss function, resulting in a model that is not only data-driven but also physically consistent. Specifically, we incorporated physics-guided constraints into the neural network by: 1) adding a physics-based term to the loss function; and 2) modifying the network architecture to reflect the underlying physical processes.

Our proposed network topology follows that of the WCM, which is widely used to model NRCS over crop fields. Through data-driven training, the model learns behaviors that are not reflected in a standard analytical WCM formulation. At the same time, the model inherits some simplifications embedded in the WCM. For example, the WCM does not represent double bounce (e.g., stem-ground) scattering. It is therefore possible that a network topology inspired by more sophisticated physical models would produce better results. However, a more complex topology may gravitate towards a fully connected dense NN, which as our results showed, is harder to train and generalizes worse.

For our physics-guided loss function, we added a single constraint, requiring the partial derivative of the NRCS with respect to the soil moisture to be positive. This constraint makes sense from a physical modeling point of view and is consistent with the WCM topology. It is also a very simple constraint, and therefore easy to implement, which we assume to be valid at all times. However, in reality, there can be a correlation between the  $SM_S$  and the VWC which can affect the NRCS in complex ways. More importantly, there is additional physical knowledge that could be incorporated into the loss function.

Back to our key hypotheses, according to our results, we can see the following.

- 1) In line with expectations, the WCM-inspired model does perform worse than the regular dense network when trained and tested on data from the same year. The WCM-inspired limits what the model can learn, e.g., hidden correlations between  $SM_S$  and VWC. However, contrary to our expectations, but in line with the literature, the results produced by the constrained models have a lower MSE for both training and test datasets. This suggests that these constraints not only prevent overfitting but also aid the training process. Our interpretation is that constraints can effectively reshape the loss landscape, potentially smoothing out poor local minima and leading the optimizer toward more generalizable and accurate solutions. Thus, when constraints align well with the underlying physical or statistical realities of the data, they can be a powerful tool to enhance the learning process and overall model performance.

- 2) As expected, both proposed approaches produce models that generalize better to completely unseen situations.
- 3) With synthetic data, the results are encouraging; the simplicity of the WCM model suggests that the neural network should learn it perfectly, yet achieving this was nontrivial and required considerable effort. This success highlights the capabilities of the network. With actual data, the intermediate outputs appear reasonable and generally consistent with typical WCM behavior. In particular, the outputs in the constrained case are quantitatively better, indicating a potential advantage of incorporating constraints.

Looking ahead, one promising direction for further research is the application of transfer learning techniques, which could leverage pretrained models to improve performance on new datasets or under different environmental conditions. In addition, addressing the challenge of local minima in neural network training remains a critical area of focus. Various strategies, including fine-tuning the network architecture and optimizing hyperparameters, may mitigate these issues and lead to more optimal solutions.

We also recommend that future studies research expand this work by extending the area of interest to cover a broader range of agricultural landscapes and by including multiple seasons to better capture seasonal dynamics. Furthermore, future efforts could focus on integrating the developed method in data assimilation or anomaly detection to enhance the practical applications of the model in agricultural monitoring. If possible, simulate the model under controlled conditions in which the expected behavior of the soil and vegetation components is known. This can help identify whether the model systematically biases towards one component or another.

#### ACKNOWLEDGMENT

The authors would like to thank our colleague at TUDelft, Dr. Riccardo Taormina, for his valuable insights into this work.

#### REFERENCES

- [1] B. A. Forman, R. H. Reichle, and C. Derksen, "Estimating passive microwave brightness temperature over snow-covered land in North America using a land surface model and an artificial neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 235–248, Jan. 2014.
- [2] B. A. Forman and R. H. Reichle, "Using a support vector machine and a land surface model to estimate large-scale passive microwave brightness temperatures over snow-covered land in North America," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 9, pp. 4431–4441, Sep. 2015.
- [3] M. Reichstein et al., "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, Feb. 2019.
- [4] S. Zhong et al., "Machine learning: New ideas and tools in environmental science and engineering," *Environ. Sci. Technol.*, vol. 55, no. 19, pp. 12741–12754, Oct. 2021.
- [5] R. Torres et al., "GMES sentinel-1 mission," *Remote Sens. Environ.*, vol. 120, pp. 9–24, May 2012.
- [6] J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, "Integrating physics-based modeling with machine learning: A survey," Mar. 2020.
- [7] M. J. Zideh, P. Chatterjee, and A. K. Srivastava, "Physics-informed machine learning for data anomaly detection, classification, localization, and mitigation: A review, challenges, and path forward," *IEEE Access*, vol. 12, pp. 4597–4617, 2024.
- [8] C. P. Murphy and J. P. Kerekes, "Physics-guided neural network for predicting chemical signatures," *Appl. Opt.*, vol. 60, no. 11, pp. 3176–3181, Apr. 2021.
- [9] A. Daw, A. Karpatne, W. Watkins, J. Read, and V. Kumar, "Physics-guided neural networks (PGNN): An application in lake temperature modeling," Sep. 2021.
- [10] X. Jia et al., "Physics-guided machine learning for scientific discovery: An application in simulating lake temperature profiles," *ACM/IMS Trans. Data Sci.*, vol. 2, no. 3, pp. 20:1–20:26, May 2021.
- [11] L. Zhong, H. Lei, and B. Gao, "Developing a physics-informed deep learning model to simulate runoff response to climate change in alpine catchments," *Water Resour. Res.*, vol. 59, no. 6, 2023, Art. no. e2022WR034118.
- [12] A. Daw, R. Q. Thomas, C. C. Carey, J. S. Read, A. P. Appling, and A. Karpatne, "Physics-guided architecture (PGA) of neural networks for quantifying uncertainty in lake temperature modeling," in *Proc. 2020 SIAM Int. Conf. Data Mining* (ser. Proceedings. Society for Industrial and Applied Mathematics), Jan. 2020, pp. 532–540.
- [13] F. Djeumou, C. Neary, E. Goubault, S. Putot, and U. Topcu, "Neural networks with physics-informed architectures and constraints for dynamical systems modeling," in *Proc. 4th Annu. Learn. Dyn. Control Conf.*, PMLR, May 2022, pp. 263–277.
- [14] H. Lievens et al., "Optimization of a radiative transfer forward operator for simulating SMOS brightness temperatures over the upper Mississippi basin," *J. Hydrometeorol.*, vol. 16, no. 3, pp. 1109–1134, Jun. 2015.
- [15] H. Lievens, B. Martens, N. E. C. Verhoest, S. Hahn, R. H. Reichle, and D. G. Miralles, "Assimilation of global radar backscatter and radiometer brightness temperature observations to improve soil moisture and land evaporation estimates," *Remote Sens. Environ.*, vol. 189, pp. 194–210, Feb. 2017.
- [16] E. P. W. Attema and F. T. Ulaby, "Vegetation modeled as a water cloud," *Radio Sci.*, vol. 13, no. 2, pp. 357–364, Mar. 1978.
- [17] M. Bracaglia, P. Ferrazzoli, and L. Guerriero, "A fully polarimetric multiple scattering model for crops," *Remote Sens. Environ.*, vol. 54, no. 3, pp. 170–179, Dec. 1995.
- [18] F. T. Ulaby, K. Sarabandi, K. McDonald, M. Whitt, and M. C. Dobson, "Michigan microwave canopy scattering model," *Int. J. Remote Sens.*, vol. 11, no. 7, pp. 1223–1253, Jul. 1990.
- [19] D. Rains, H. Lievens, G. J. M. De Lannoy, M. F. McCabe, R. A. M. de Jeu, and D. G. Miralles, "Sentinel-1 backscatter assimilation using support vector regression or the water cloud model at European soil moisture sites," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2021.
- [20] X. Shan et al., "Towards constraining soil and vegetation dynamics in land surface models: Modeling ASCAT backscatter incidence-angle dependence with a deep neural network," *Remote Sens. Environ.*, vol. 279, Sep. 2022, Art. no. 113116.
- [21] S. de Roos, L. Busschaert, H. Lievens, M. Bechtold, and G. J. M. De Lannoy, "Optimisation of AquaCrop backscatter simulations using Sentinel-1 observations," *Remote Sens. Environ.*, vol. 294, Aug. 2023, Art. no. 113621.
- [22] T. Nikaiein, P. Lopez-Dekker, S. C. Steele-Dunne, V. Kumar, and M. Huber, "Modeling SAR observables by combining a crop-growth model with machine learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 7763–7776, 2023.
- [23] C. A. Jones and J. R. Kiniry, *A CERES-Maize: Simulation Model of Maize Growth and Development*. Austin, TX, USA: Texas A&M University Press, 1986.
- [24] J. W. Jones et al., "Decision support system for agrotechnology transfer: DSSAT V3," in *Understanding Options for Agricultural Production* (ser. Systems Approaches for Sustainable Agricultural Development), G. Y. Tsuiji, G. Hoogenboom, and P. K. Thornton, Eds. Dordrecht, Netherlands: Springer, 1998, pp. 157–177.
- [25] V. Kumar, M. Huber, B. Rommen, and S. C. Steele-Dunne, "Agricultural SandboxNL: A national-scale database of parcel-level processed Sentinel-1 SAR data," *Sci. Data*, vol. 9, no. 1, Jul. 2022, Art. no. 402.
- [26] M. C. Dobson and F. T. Ulaby, "Active microwave soil moisture research," *IEEE Trans. Geosci. Remote Sens.*, vol. GE-24, no. 1, pp. 23–36, Jan. 1986.
- [27] K. Chen, T.-D. Wu, L. Tsang, Q. Li, J. Shi, and A. Fung, "Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 1, pp. 90–101, Jan. 2003.
- [28] S. C. Steele-Dunne, H. McNairn, A. Monsivais-Huertero, J. Judge, P.-W. Liu, and K. Papathanassiou, "Radar remote sensing of agricultural canopies: A review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 2249–2273, May 2017.

- [29] I. Champion and G. Guyot, "Generalized formulation for semi-empirical radar models representing crop backscattering," 1991.
- [30] L. Prévot, I. Champion, and G. Guyot, "Estimating surface soil moisture and leaf area index of a wheat canopy using a dual-frequency (C and X bands) scatterometer," *Remote Sens. Environ.*, vol. 46, no. 3, pp. 331–339, Dec. 1993.
- [31] F. Chollet, "Keras," 2015. [Online]. Available: <https://github.com/fchollet/keras>
- [32] M. Shorachi, V. Kumar, and S. C. Steele-Dunne, "Sentinel-1 SAR backscatter response to agricultural drought in The Netherlands," *Remote Sens.*, vol. 14, no. 10, Jan. 2022, Art. no. 2435.



**Tina Nikaein** (Member, IEEE) received the B.Sc. degree in surveying and geomatics engineering and the M.Sc. degree in remote sensing from Tehran University, Tehran, Iran, in 2015 and 2018, respectively, and the Ph.D. degree in remote sensing from the Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands.

Her research interests include radar remote sensing, InSAR, and environmental monitoring.



**Paco Lopez-Dekker** (Senior Member, IEEE) was born in Nijmegen, The Netherlands, in 1972. He received the Ingeniero degree in telecommunication engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1997, the M.S. degree in electrical and computer engineering from the University of California, Irvine, CA, USA, in 1998, under the Balsells Fellowship, and the Ph.D. degree in clear-air imaging radar systems to study the atmospheric boundary layer from the University of Massachusetts, Amherst, MA, USA, in 2003.

In 2003, he joined Starlab Barcelona, where he worked on the development of GNSS-R sensors and techniques. From 2004 to 2006, he was a Visiting Professor with the Department of Telecommunications and Systems Engineering, Universitat Autònoma de Barcelona. In March 2006, he was awarded a Ramon y Cajal Grant to conduct pioneering research on bistatic synthetic aperture radar (SAR) at Remote Sensing Laboratory, UPC. At the university, he taught courses on signals and systems, signal processing, communications systems and radiation, and guided waves. From 2009 to 2016, he led the SAR Missions Group at the Microwaves and Radar Institute, German Aerospace Center, Wessling, Germany. Since 2016, he has been an Associate Professor with the Department of Geoscience and Remote Sensing, Faculty of Civil Engineering and Geosciences. He has been deeply involved in several radar mission proposals, and is the Lead Investigator of the Harmony ESA Earth Explorer 10 Mission candidate. He has authored or coauthored 50 peer-reviewed journal articles and more than 100 conference contributions in a broad range of topics related to radar remote sensing.