FORECASTING SOIL MOISTURE USING A DEEP LEARNING MODEL INTEGRATED WITH PASSIVE MICROWAVE RETRIEVAL

Archana Kannan[1], Grigorios Tsagkatakis[2], Ruzbeh Akbar[3], Daniel Selva[4], Vinay Ravindra[5,6], Richard Levinson[5,7], Sreeja Nag[5,6], Mahta Moghaddam[1]

[1]University of Southern California, Los Angeles, CA, United States, [2]Foundation for Research and Technology - Hellas (FORTH), Institute of Computer Science, Heraklion, Greece, [3]Massachusetts Institute of Technology, Cambridge, MA, United States, [4]Texas A&M University, College Station, TX, United States, [5]NASA Ames Research Center, Moffett Field, CA, United States, [6]Bay Area Environmental Research Institute Moffett Field, Moffett Field, CA, United States, [7]KBR Wyle Services, LLC, Torrance, CA, United States

ABSTRACT

In this paper we develop a Convolutional Long Short-term memory (ConvLSTM) model, a time series deep learning neural network, to predict soil moisture, with an add-on module of passive microwave (radiometer) soil moisture retrieval using the Tau omega model. We incorporate antecedent observations, landscape properties, and forcing factors such as precipitation, landcover, clay fraction, and brightness temperature in the prediction scheme. A regularization Monte Carlo Dropout layer is added to the network to remove stochasticity and avoid overfitting during the training phase. This dropout layer also provides a Bayesian approximation to quantify uncertainty during forecasting. The model is validated at four Soil Moisture Active Passive (SMAP) Cal/Val locations using performance metrics such as Root Mean Square Error (RMSE) and Bias to evaluate effectiveness of the proposed method. This model is developed as a component of the Science Simulator within the Distributed Spacecraft with Heuristic Intelligence to Enable Logistical Decisions (D-SHIELD) project.

Index Terms— ConvLSTM, Tau-Omega model, Soil Moisture

1. INTRODUCTION

Soil Moisture has complex structural characteristics and is influenced by various spatial, temporal, and meteorological conditions. Developing a mathematical model to capture this high degree of spatiotemporal heterogeneity in predicting soil moisture fields is a challenging task.

Many methods have been proposed in the literature to estimate soil moisture. With the recent advancements in Machine Learning (ML) and availability of satellite based remote sensing observations, data-driven models have been developed to estimate soil moisture. Cai et al. investigated soil moisture prediction using a deep learning regression network in the Beijing area in China [1]. Fang and Shen produced a near real time forecast of soil moisture using an LSTM model [2] for the CONUS area.

In this work, we adapt the ConvLSTM [3] deep learning network to forecast global surface soil moisture three days into the future. The model predicts soil moisture along with an uncertainty layer, which is determined using the added regularization Monte Carlo dropout layer [4], [5]. This predicted soil moisture and uncertainty are gridded at $9km \times 9km$ resolution. The pixels (spatial location) or Ground Points (GPs) with high uncertainty are then assimilated with the simulated soil moisture from the passive microwave (radiometer) retrieval process using the Tau-omega model [6]. Both assimilated and unassimilated predictions are compared with the NASA SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data (SPL4SMGP) product [9] to assess the accuracy of the predictions.

This model is developed as a component of the D-SHIELD project Science Simulator [7]. D-SHIELD consists of software tools designed to plan and schedule spacecraft payloads and operations, to improve global surface soil moisture monitoring via various microwave remote sensing assets. The Simulator predicts surface soil moisture and its prediction uncertainty, within a finite, but variable, prediction window, which enables D-SHILED constellation planner and scheduler to determine optimum payload and instrument configurations for soil moisture observations.

2. SOIL MOISTURE FORECAST MODEL

The forecast process comprises of two steps: ConvLSTM deep learning network and Tau-omega physics-based radiometer observation model. For training the ConvLSTM model, SMAP L4 geophysical products (soil moisture, precipitation) are used. For simulating passive microwave products, the brightness temperature (T_B) is derived from 36 km SMAP footprint using Backus-Gilbert interpolation on

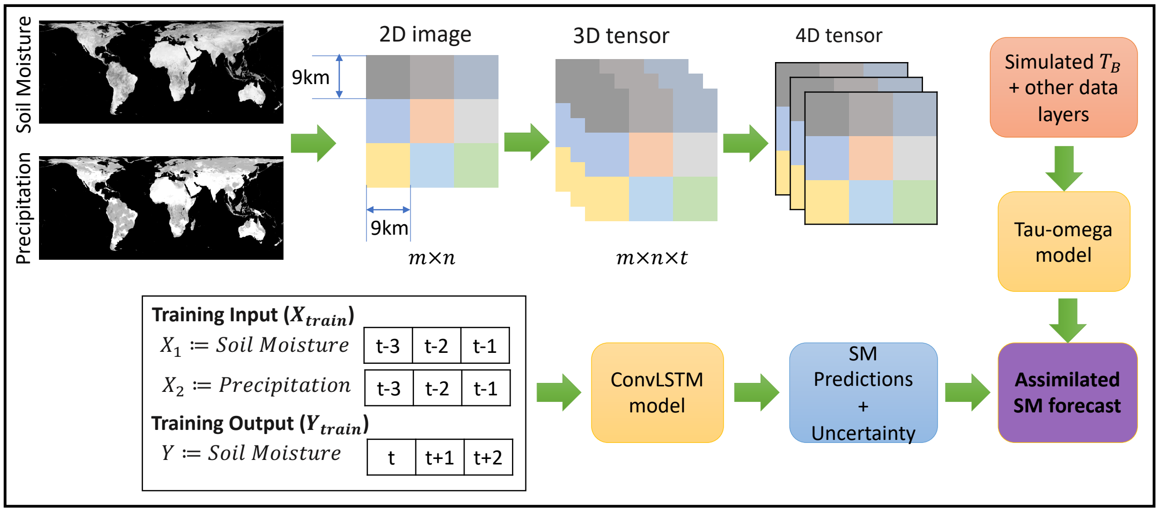


Figure 1. Framework for forecasting global soil moisture using SMAP satellite driven geophysical data products

the 9 km EASE-Grid using the radiometer vertical polarization.

Using T_B along with surface temperature, vegetation water content, and clay percent in the soil, the values of soil moisture are estimated using the Tau-omega model. Figure 1 explains the forecasting process in detail.

2.1 ConvLSTM deep learning model

ConvLSTM is gaining popularity in various fields for its ability to predict future state using a series of past images. The future state of each pixel is determined by the input and past states of its local neighbors. The neural network trained to predict soil moisture comprises of convolutional layers that perform convolution manipulations on the input images and intermediate layers; concatenation layer for combing antecedent soil moisture, precipitation images; and a dropout layer.

The model is trained with the 2015-2019 historical SMAP L4 soil moisture and precipitation data with a chosen window size. This window size determines the temporal resolution of predictions. Once the model is trained, it is used to predict soil moisture for the year 2020, with SMAP L4 soil moisture as input, three days into the future. The prediction is repeated for multiple times to calculate mean prediction and standard deviation from the mean, which is considered as the prediction uncertainty.

2.2 Tau-omega model

The GPs, where the uncertainty of soil moisture prediction from ConvLSTM model is higher than 0.04 m^3/m^3, is assimilated with the radiometer retrieved soil moisture. The baseline algorithm for passive microwave soil moisture retrieval follows the SMAP Single Channel Algorithm at Vertical polarization [6].

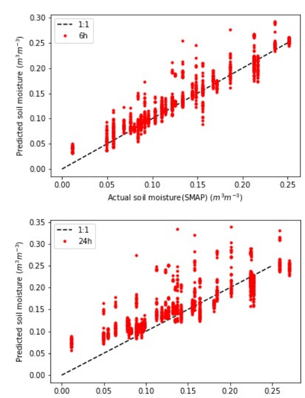
Using the following data layers: vegetation parameter depending on the landcover type (b), surface temperature (T_s), vegetation water content (VWC), single scattering albedo (ω), incidence angle (θ), the transmissivity (γ) is calculated by $\gamma = exp(-(b \times VWC)\, sec\theta)$. Using γ, reflectivity (r) is calculated by:

$$r = \frac{T_s(1-\omega+\omega\gamma) - T_B}{T_s\gamma(\gamma+\omega-\omega\gamma)}$$

From reflectivity, using Fresnel equation inversions, dielectric constant (ε) is computed as follows:

$$\varepsilon = \frac{sin^2\theta}{cos\theta}\left[\frac{(1+r)cos\theta}{(1+r)-(1-r)cos\theta}\right]$$

Finally, the calculated dielectric constant along with clay percent is used in the Mironov dielectric model [8] to estimate soil moisture.



Interval	RMSE	Bias
6 hours	0.0213	0.00178
12 hours	0.0324	0.00217
24 hours	0.0383	0.00284

Figure 2. Comparison of predicted and SMAP SM for different prediction windows at Tonzi ranch, CA, USA

3. CHOOSING PREDICTION WINDOW

The prediction temporal resolution is dependent on the tensor size of the inputs to the ConvLSTM model. Different tensor sizes within the model were explored. A localized experiment was performed on an area of $54km \times 54km$ grid centered at 30.3°N, 120.9°W at Tonzi Ranch CA, USA. SMAP L4 product is available in 3-hourly resolution. Combining 3, 6, and 8 datapoints, 6-, 12-, 24- hours predictions were generated. The results are shown in Figure 2. The 6-hour prediction window exhibited lowest error of 0.02 m^3/m^3. From this, the 6-hour window option was chosen for global predictions of soil moisture.

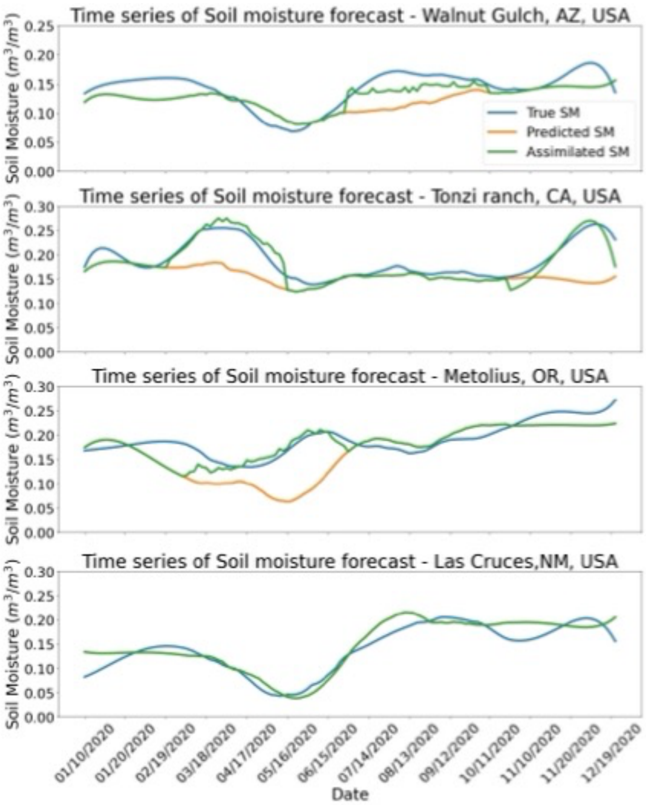


Figure 3. Soil moisture daily average forecasts for (a) Walnut gulch, (b) Tonzi ranch, (c) Metolius, (d) Las Cruces for a year

4. EXPERIMENTAL RESULTS

The forecast time series for four SMAP Cal/Val locations is given in Figure 3 and the performance metrics are in Table 1. The prediction varies with different landcover type. For dry regions the prediction follows true soil moisture with no need for assimilation as seen in location (d). Precipitation and drying down is not entirely captured by the deep learning model, leading to higher uncertainty and the need for assimilation with simulated soil moisture.

The ConvLSTM model exhibits $0.01 - 0.05\, m^3/m^3$ RMSE globally. With addition of radiometer retrievals there is an error reduction of 20-50%.

5. CONCLUSION

In this work we developed a time series deep leaning network to estimate soil moisture from satellite derived data products with a radiometer retrieval module to assimilate high uncertainty predictions. This predictive model is part of the Science Simulator module within the D-SHIELD project. In future work, including radar and other microwave sensing resources to the predictions will be explored.

(NOTE: All deep leaning simulations were done using Google TensorFlow Keras library.)

Table 1. Performance metrics for different locations in Figure 4. Pred - Predicted SM; Assim - Assimilated SM with retrieved SM

Loc	IGBP class	RMSE		Bias	
		Pred	Assim	Pred	Assim
(a)	Shrubland	0.0293	0.0208	0.0194	0.0127
(b)	Woody savanna	0.0466	0.0149	0.0319	0.0051
(c)	Evergreen forest	0.0457	0.0230	0.0248	0.0055
(d)	Bare surface	0.0192	0.0192	0.0070	0.0070

6. REFERENCES

[1] Cai Y, Zheng W, Zhang X, Zhangzhong L, Xue X. “Research on soil moisture prediction model based on deep learning”. PLoS ONE. 2019;14(4):1–19.

[2] Fang K, Shen C. “Near-real-time forecast of satellite-based soil moisture using long short-term memory with an adaptive data integration kernel”. Journal of Hydrometeorology. 2020;21(3):399–413.

[3] Chao Z, Pu F, Yin Y, Han B, Chen X. “Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting”. Journal of Sensors. 2018; 2018:1–9.

[4] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”. Journal of Machine Learning Research. 2014; 15:1929–58.

[5] Gal Y, Ghahramani Z. “Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning”. Proceedings of the 33rd International Conference on Machine Learning, New York, NY, USA, 2016 JMLR: W&CP. 2016;48:1022–34.

[6] Entekhabi D, Das N, Njoku E, Yueh S, Johnson J, Shi J. “Soil Moisture Active Passive (SMAP) Algorithm Theoretical Basis Document L2 & L3 Radar/Radiometer Soil Moisture (Passive) Data Products”. JPL, 2014;1–89.

[7] Nag S, Moghaddam M, Selva D, Frank J, Ravindra V, Levinson R, Azemati A, Aguilar A, Li A, Akbar R, “D-SHIELD: Distributed Spacecraft with Heuristic Intelligence to Enable Logistical Decisions”. IEEE International Geoscience and Remote Sensing Symposium (*IGARSS* 2020)

[8] Mironov VL, Kosolapova LG, Fomin SV. “Physically and mineralogically based spectroscopic dielectric model for moist soils”. IEEE Transactions on Geoscience and Remote Sensing. 2009;47(7):2059–70.

[9] Reichle, R., G. De Lannoy, R. D. Koster, W. T. Crow, J. S. Kimball, and Q. Liu. 2020. SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Analysis Update, Version 5. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: https://doi.org/10.5067/0D8JT6S27BS9.