# Calibration of Dobson model for improving soil moisture retrievals from AMSR satellite imagery

https://doi.org/10.31713/MCIT.2023.005

Mykhailo Boiko
EOS Data Analytics,
Department of Computer Sciences and Applied Mathematics
National University of Water and Environmental Engineering
Rivne, Ukraine
m.v.boiko@nuwm.edu.ua

Olena Belozerova
EOS Data Analytics,
Department of Computer Sciences and Applied Mathematics
National University of Water and Environmental Engineering
Rivne, Ukraine

*Abstract*—**Satellite-based retrieval of environmental variables has seen rapid advancements in recent years, although it remains challenged by various sources of uncertainty. In this study, we endeavor to enhance the accuracy of soil moisture estimation from AMSR imagery through parameter calibration. Our focus is on calibrating the Dobson dielectric mixing model, a component of the radiative transfer model that relies heavily on empirical relationships.**

**The calibration process is based on black-box optimization techniques for the retrieval problem. We have adapted the CORS optimization algorithm to address the specific characteristics of our task. We also considered different target functions for calibration.**

**To evaluate the efficacy of our framework, we conducted tests across a dataset comprising 118 ground stations in the United States. The outcomes reveal that optimizing parameter settings can provide limited improvement to the accuracy and, when configured accordingly, address specific issues such as bias correction. Calibration emerges as a potent tool for refining surface soil moisture retrievals, although its effectiveness tends to diminish in larger calibration areas.**

**Keywords—satellite soil moisture; model calibration; black-box optimization; Dobson model.**

## I. Introduction

Soil moisture (SM) measurement is essential in comprehending the hydrological cycle and its implications on weather and climate. Precise estimations of soil moisture are essential for advancing our knowledge of the surface energy budget and research of hydrological, ecological and atmospheric processes.

Remote sensing satellites have emerged as a formidable tool in the hydrological community for retrieving soil moisture. Despite their capability to assess only the topmost soil layer (typically 1–5 cm), it nevertheless carries significant importance across various environmental domains, encompassing hydrology, meteorology, agriculture, and climate change.

Microwave techniques have obtained widespread recognition for their potential in routine soil moisture retrieval, whether through active or passive sensors. Notably, soil moisture retrieval from Advanced Microwave Scanning Radiometer (AMSR-E) data has exhibited lower accuracy compared to active (e.g., Sentinel-1) and active-passive sensors (SMAP). Research findings have consistently pointed to systematic biases in AMSR-E's soil moisture observations. One of the ways to enhance the quality of soil moisture retrievals is calibration of LPRM parameters [1].

In this study, we concentrate on the parameters embedded within the Dobson soil-water dielectric mixing model. This model, apart from satellite SM retrieval, finds application in other soil moisture monitoring technologies, such as in-situ soil moisture sensors. The Dobson model serves as a means to convert sensed dielectric constant values into soil moisture, and it has seen calibration adaptations to accommodate specific properties such as soil temperature [2], clay content [3], and tillage [4].

The task of calibrating LPRM requires a cost-effective optimization algorithm. Given the absence of analytical rules that can be derived the model, the task necessitates utilization of what is commonly referred to as a black-box optimization method. Such methods encompass various techniques, some of which are exemplified by [5, 6]. A common feature among them is incorporation of a surrogate – an approximated hyperplane that is subject to optimization instead on the original function. This surrogate strategy is effective for high-dimensional problems and particularly advantageous for dealing with the computational expense associated with function evaluations [7].

In our specific context, we employ a modified CORS algorithm (COordinate search using Response Surfaces) [8] to fine-tune the parameters of the Dobson model. We propose further adaptations for this method to adapt it for calibration tasks.

## II. Models and Methods

### A. *Land Parameter Retrieval Model (LPRM) and Dobson model*

Our focus in this study is on soil moisture retrieval via the Advanced Microwave Scanning Radiometer (AMSR-E), which relies on passive microwave theory and employs the Land Parameter Retrieval Model (LPRM) to facilitate this process.

Dielectric mixing models covers the relation between soil moisture and dielectric constant, that is further related to brightness temperature. It takes advantage of the substantial disparity in dielectric constants of dry soil ($\varepsilon \approx 6$) and water ($\varepsilon \approx 80$) at microwave frequencies [9].

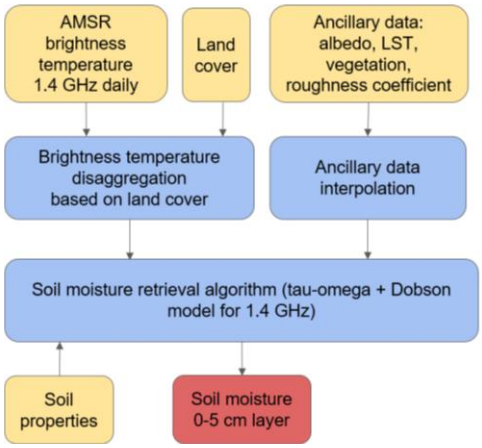



Figure 1. Schematic representation of the key steps in the Land Parameter Retrieval Model (LPRM). Soil moisture is calculated by matching the observed brightness temperature with the model-predicted brightness temperature from radiative transfer

One of the most popular dielectric mixing models is a semi-empirical model developed by Dobson et al. [10]. It covers a broad frequency range, between 1.4 and 18 GHz, and provides both the real and imaginary components of the dielectric constant in terms of the soil texture (% sand, silt and clay), bulk density and volumetric soil moisture [9]. According to the semi-empirical formulation of Dobson model, complex dielectric constant $\varepsilon$ is calculated as

$$\begin{aligned} &Re(\varepsilon) = \left(1 + 0.66\rho_b + \theta^{\beta_1} \cdot Re(\varepsilon_w)^{\alpha}\right)^{\frac{1}{\alpha}}, \\ &Img(\varepsilon) = \theta^{\beta_2} \cdot Img(\varepsilon_w), \end{aligned} \qquad (1)$$

where $\theta$ is volumetric soil moisture, $\varepsilon_w$ is complex dielectric constant of the water contained in the soil, $\rho_b$ – soil bulk density. Coefficients $\alpha, \beta_1, \beta_1$ in (1) are empirical parameters estimated as follows (according to [11]):

$$\begin{aligned} &\alpha = 0.65, \\ &\beta_1 = 1.27 - 0.519\, S - 0.152\, C, \\ &\beta_2 = 2.06 - 0.928\, S - 0.255\, C. \end{aligned} \qquad (2)$$

The key factor determining dielectric constant of water is soil temperature. They are linked through the static dielectric constant of water $e_{w0}$, empirically calculated as

$$\begin{aligned} &e_{w0} = 88.045 - 0.4147\, T + 6.295 \cdot 10^4\, T^2 + \\ &+1.075 \cdot 10^5\, T^3, \end{aligned} \qquad (3)$$

where T is soil temperature.

Relations (2), (3) are deciding in estimating soil dielectric constant. Therefore, we selected the empirical coefficients in these relations as target for calibration. Empirical parameters $\beta$ are the most obvious candidates for calibration, since they determine the model response to the soil composition. Moreover, we add soil bulk density $\rho_b$ as another important soil parameter that is already in use by the model. The changes to parameter $\alpha$ are suggested in a number of studies, including [2]. As for the temperature dependent dielectric constant of water $e_{w0}$, more accurate temperature response might help to adapt the model to actual climatic conditions. The chosen calibration ranges for parameters are given in Table 1.

TABLE I. CHOSEN CONSTRAINTS FOR CALIBRATION OF THE DOBSON MODEL PARAMETERS

| **Parameter** | **Term** | **Default value** | **Calibration constraints** | |
|---|---|---|---|---|
| | | | ***Minimum*** | ***Maximum*** |
| alpha | 1 | 0.65 | 0.6 | 0.7 |
| beta1 | 1 | 1.270 | 1.143 | 1.397 |
| | sand | -0.519 | -0.571 | -0.467 |
| | clay | -0.152 | -0.167 | -0.137 |
| | density | 0 | -0.1 | 0.1 |
| beta2 | 1 | 2.06 | 1.864 | 2.266 |
| | sand | -0.928 | -1.021 | -0.835 |
| | clay | -0.255 | -0.281 | -0.230 |
| | density | 0 | -0.1 | 0.1 |
| ew0 | 1 | 88.045 | 79.241 | 96.850 |
| | T | -0.414 | -0.456 | -0.373 |
| | T2 | 0.00063 | 0.0004 | 0.00075 |

### B. *Modified CORS calibration algorithm*

We seek to calibrate the semi-empirical Dobson model for optimal alignment with in situ soil moisture data. This calibration involves minimizing soil moisture estimation errors within the Land Parameter Retrieval Model (LPRM), which is computationally intensive and implicitly defined, necessitating the use of an efficient black-box optimization algorithm.

CORS (COordinate search using Response Surfaces) is our chosen algorithm for constrained, costly black-box optimization. It employs radial basis function (RBF) surrogates to approximate the objective function, controlling the time cost by defining a specific number of expensive function calls. This allocation is divided into two phases:

1) *Initialization Phase:* Here, the algorithm evaluates the function at randomly selected points to construct an initial surrogate function. The most promising points are then identified.
2) *Refinement Phase:* Building on the promising points, the algorithm iteratively suggests new points optimized by the surrogate hyperplane. These points are evaluated for function values and contribute to the ongoing refinement of the surrogate. This process continues until remaining resource of points is exhausted.

The necessary prerequisite for applying the above technique is to normalize the value constraints of all calibrated parameters to a range of [0; 1]. This normalization transforms the search space into a uniform unit cube, ensuring consistent distances for surrogate optimization.

The initial set of points has to be distributed evenly across the problem space. Original implementation of the algorithm implemented Latin hypercube approach for point spacing. However, we found that in case of high dimensionality the Poisson disk sampling method produces a more even distribution of points. This algorithm is notable for ensuring that points are located no closer to each other than a pre-set minimum distance, resulting in a more intuitive distribution [12].

The original algorithm also suggests that the given number of function calls should be split equally between initial and refinement points. After experimenting with the point division, we found that increasing the number of initial points above a specific number provides no additional improvement to final result. In our case, 50 points proved enough for a reasonably good initial distribution. The remaining points (25 at least) are used to refine the initial estimation.

Also, a number of other minor modifications to the algorithm proposed in [8] was introduced. We removed normalization step before estimation of target function, since SM values are already contained within the [0; 1) interval and additional scaling could lead to misinterpretations. Multiquadric radial basis function was chosen for surrogate approximation, and the initial SLSQP optimization method was replaced by Nedler-Mead.

### C. *Example test of optimization method*

Before embarking on the actual validation experiment, we conducted a preliminary assessment of our modifications using a theoretical problem. The task involved optimizing the valley function defined as

$$f(x, y) = |x - y| + \left(\frac{x + y - 1}{3}\right)^2, 0 \leq x, y \leq 1.$$

While the optimal solution of the problem is obvious analytically, functions with valley-like characteristics are well-known for posing significant optimization challenges.

In Figure 2, we present simulation details for both the original and adapted algorithms, considering a limitation of 100 points. Our modified algorithm produced a better solution by focusing on local minima points identified through the initial distribution. Conversely, the original algorithm extensively explored a more diverse set of points, leading to the inability to reach the optimal solution within the prescribed number of points.

### D. *Experimental setting*

In this study, we take disaggregated AMRS-2 L-band satellite imagery. The image data are processed with LPRM model to estimate surface SM. These results are further denoted as reference values. Both the reference and values produced after model calibration are validated against ground data using the following metrics: root mean squared error (RMSE), correlation (R), average relative error (ARE) and index of agreement (IoA).

To validate our results, we utilize in-situ sensor data sourced from the International Soil Moisture Network (ISMN) [13]. Due to technical reasons, we selected for analysis a single imagery tile in North East of the USA, comprising the area between 41.35° and 49.65° latitude and between -116.2° and -107.25° longitude, with a total of 118 ground stations. For both satellite imagery and in-situ datasets we use records for the whole length of 2015.

## III. Results

We perform calibration with different target metrics: RMSE, ARE and correlation, as well as evaluation of reference values. Summary of validation results is presented in Table 2. Here units for RMSE, ARE and bias are the same as SM units, $m^3 \cdot m^{-3}$.

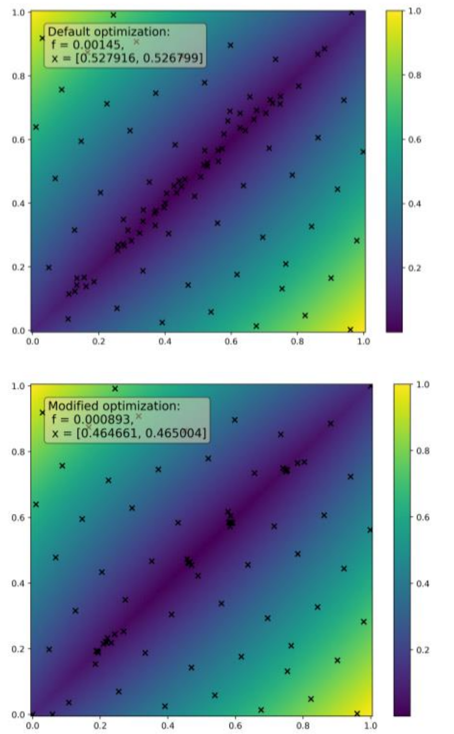



Figure 2. Performance of the original (top) and modified (bottom) optimization algorithms for a valley function

The metrics presented in Table 2 indicate that calibration can only achieve partial improvement. Correlation between the satellite retrievals and ground data is very low in the selected region due to the terrain complexity and presence of different soil types. Moreover, results for ARE metric are extremely high due to the fact that SM values can be close to 0.

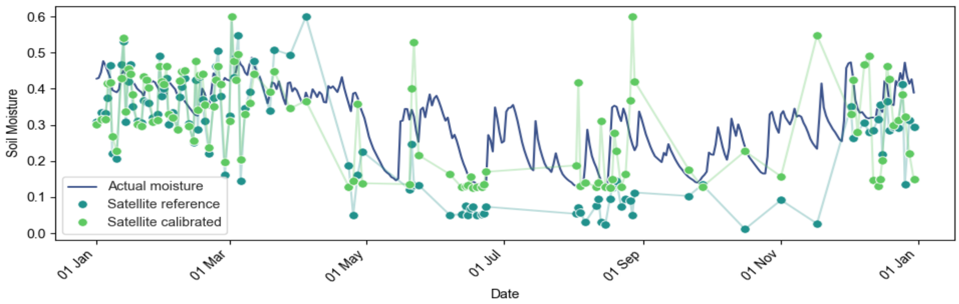
The results vary greatly with chosen calibration metric. Calibration improves result for the target metric, yet this is achieved at expense of other characteristics. Therefore, the choice of target metric must be made according to the desired properties of the result. RMSE is the most stable metric for removing bias and variation errors. Correlation should be chosen if general agreement of trends is the priority. Surprisingly, ARE proved to be a good conservative metric that provides improvement without loss of agreement.

TABLE II. Validation results for SM satellite estimates, including reference values and calibrated by RMSE, ARE and correlation

| **Calibration metric** | **RMSE** | **ARE** | **R** | **IoA** |
|---|---|---|---|---|
| Reference | 0.157 | 1.152 | 0.087 | 0.436 |
| RMSE | 0.114 | 4.979 | -0.110 | 0.244 |
| ARE | 0.153 | 0.674 | -0.103 | 0.436 |
| R | 0.171 | 1.267 | 0.126 | 0.449 |

Further, we investigate the influence of calibration of the final result. For this purpose, we select one of the ground stations from the dataset and compare the estimated SM values before and after calibration. Fig. 3 shows the time series of actual, reference and calibrated SM on station McAlister Farm.

The moisture trends on Fig. 3 indicate that calibration is capable of shifting the estimates in a way that is close to linear scaling, with occasional deviations.




Comparison of reference and calibrated satellite estimates with in-situ soil moisture data on station McAlister Farm

The initial reference estimate underestimated actual soil moisture, but calibrated version attempts to remove this obvious bias. However, the estimated trend remains largely unchanged. This example agrees with our previous results in the fact that correlation cannot be greatly improved with parameter calibration. Instead, target metrics that compare actual values (such as RMSE and ARE) can decrease bias and yield overall better results.

## IV. Conclusion

In this study, we calibrated Dobson model to improve soil moisture estimation from AMSR satellite data. We applied a surrogate-based optimization algorithm to determine optimal parameters of the model. We also proposed modifications to the CORS optimization algorithm, which have demonstrated their effectiveness for the functions of high complexity with multiple local minima.

The findings strongly suggest that calibrating components within the Land Parameter Retrieval Model presents an effective approach for mitigating bias in soil moisture estimates. This calibration method is most effective at the local or regional level.

The choice of an appropriate target function for calibration emerges as a critical factor in achieving specific optimization objectives, whether it involves enhancing agreement or removing bias. This flexibility enables tailored results aligned with specific research goals.

## Acknowledgment

We extend our sincere appreciation to EOS Data Analytics for their generous support in conducting this research, providing data and computational resources.

## References


[1] D. Al-Sharafany, “Soil moisture retrieval from the AMSR-E,” Agricultural Water Management. Theories and Practices, London: Academic Press, 2021, pp. 241–278.

[2] M. Schwartz, Z. Li, T. Sakaki, A. Moradi, and K. Smits, “Accounting for Temperature Effects on the Performance of Soil Moisture Sensors in Sandy Soils,” Soil Sci. Soc. Am. J., vol. 83, 2019, pp. 1319–1323.

[3] R.C. Schwartz, S.R. Evett, and J.M. Bell, “Complex Permittivity Model for Time Domain Reflectometry Soil Water Content Sensing: II. Calibration,” Soil Sci. Soc. Am. J., vol. 73, 2009, pp. 898–909.

[4] R.C. Schwartz, A.J. Schlegel, J.M. Bell, R.L. Baumhardt, and S.R. Evett, “Contrasting tillage effects on stored soil water, infiltration and evapotranspiration fluxes in a dryland rotation at two locations,” Soil and Tillage Res., vol. 190, 2019, pp. 157–174

[5] P.S. Palar, Y.B. Dwianto, R.G. Regis et al., “Benchmarking Constrained Surrogatebased Optimization on Low Speed Airfoil Design Problems,” Proceedings of Genetic and Evolutionary Computation Conference (GECCO’19, July 13–17, 2019), Companion, Prague, Czech Republic, 2019, pp. 1990–1998.

[6] P. Beaucaire, Ch. Beauthier, and C. Sainvitu, “Multi-point infill sampling strategies exploiting multiple surrogate models,” Proceedings of Genetic and Evolutionary Computation Conference (GECCO’19, July 13–17, 2019), Companion, Prague, Czech Republic, 2019, 1559–1567.

[7] R.G. Regis, “A Survey of Surrogate Approaches for Expensive Constrained Black-Box Optimization,” in Optimization of Complex Systems: Theory, Models, Algorithms and Applications, H.L. Thi, H. Le and T.P. Dinh, Eds. Cham: Springer, 2019, pp. 37–47.

[8] P. Knysh and Y. Krokolis. A procedure for parallel optimization of expensive black-box functions. [Online]. Available: https://arxiv.org/pdf/1605.00998.pdf. [Accessed: Octember 15, 2023].

[9] B. W. Barrett, E. Dwyer, and P. Whelan, “Soil Moisture Retrieval from Active Spaceborne Microwave Observations: An Evaluation of Current Techniques,” Remote Sens., vol. 1, 2009, pp. 210–242.

[10] M.C. Dobson, F.T. Ulaby, M.T. Hallikainen, and M.A. El-Rayes, “Microwave Dielectric Behavior of Wet Soil-Part II: Dielectric Mixing Models,” IEEE Trans. Geosc. Remote Sens., vol. GE-23, pp. 35-46, January 1985.

[11] F.T. Ulaby, D.G. Long, and W.J. Blackwell et al., “Microwave Radar and Radiometric Remote Sensing,” Ann Arbor: University of Michigan Press, 2013.

[12] R. Bridson, “Fast Poisson disk sampling in arbitrary dimensions,” In ACM SIGGRAPH 2007 sketches (SIGGRAPH '07), Association for Computing Machinery, New York, USA, pp. 22-es, 2007.

[13] W. Dorigo, W. Wagner, R. Hohensinn 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, 2011, pp. 1675–1698.