

This is a review of the manuscript “Interpretable Soil Moisture Prediction with a Physics-Guided Deep Learning Approach.” The authors propose non-local neural networks (NLNNs) for single-time-step, multi-depth soil-moisture forecasts, with two variants: a self-attention NLNN (SA-NLNN) and a physics-guided NLNN (PG-NLNN) that disentangles four influences (upper boundary, upper layers, same-depth memory, lower layers) motivated by gravity, capillarity, and retention. They test their models on both synthetic and field data. They compare the performance of the two models with LSTM baselines and show that the prediction uncertainty is smaller for the proposed NN models than for the LSTM models. They also show that the learned non-local weight matrices can be related to soil texture. This is an interesting direction of research.

I believe this manuscript contains many good ideas for further investigation. Thus, I encourage publication after a major revision.

Response:

We sincerely appreciate the reviewer’s time and effort in evaluating our manuscript and providing insightful comments. The comments have significantly improved the quality of our work.

Since our model is governed not by explicit physical equations but by fundamental knowledge of soil water dynamics, we have replaced the term "physics-guided" with the more accurate descriptor "knowledge-guided" throughout the manuscript to better reflect this approach.

Major points

Synthetic vs. field data

It would be more natural to use the same set of models for both synthetic and field data. LSTM baselines are evaluated on field data; the synthetic section compares SA-NLNN vs. PG-NLNN but omits LSTM models on the same synthetic tasks. This weakens causal attribution of PG-NLNN’s gains to physics guidance rather than dataset characteristics.

Response1:

Thank you for your comments. We have provided the performance of the LSTM models on four synthetic cases and compared them with our proposed models. Table R1 displays the MAE values for 1, 3, 7, and 15-day forecasts of four models. Figure R1 visually displays their RMSE results for both heterogeneous and layered soil scenarios. As shown in Table R1, the LSTM_4 model achieves very high accuracy in 1-day predictions, but its performance deteriorates rapidly over longer periods. As for the other models, NLNNs and LSTM_1 exhibit comparable performance, with LSTM_1 in fact outperforms them in certain scenarios, for example, in the case of root water uptake. To further investigate and compare these four models, we evaluate them under five levels of noise (0.5, 1.0, 2.0, 5.0, 10.0). The RMSE results for soil moisture prediction at 0.05m, 0.10m, 0.20m, 0.50m, and 1.00m are presented in Figure R2. The LSTM_4 model demonstrates poor noise resistance and long-term forecasting capability. The other three models perform similarly under low-noise conditions, with LSTM_1 even exhibiting some advantage. However, as the noise level increases, the two NLNN models demonstrate better robustness. Notably, the knowledge-guided NLNN is particularly stable, consistent with its performance on in-situ soil moisture data.

Tabel R1. The MAE [%] values for 1, 3, 7, and 15-day forecasts of LSTM_4, LSTM_1, the proposed KG-NLNN and SA-NLNN model at 5 depths under four designed scenarios.

Depth/m	KG-NLNN															
	homogeneous				heterogeneous				two-layered				root water uptake			
	1d	3d	7d	15d	1d	3d	7d	15d	1d	3d	7d	15d	1d	3d	7d	15d
0.05	0.235	0.327	0.433	0.539	0.259	0.372	0.510	0.652	0.449	0.680	0.945	1.170	0.528	0.747	0.996	1.224
0.10	0.313	0.451	0.627	0.788	0.306	0.431	0.593	0.749	0.521	0.745	0.995	1.191	0.409	0.544	0.685	0.825
0.20	0.342	0.533	0.776	1.016	0.305	0.488	0.736	0.971	0.433	0.649	0.901	1.179	0.623	0.852	1.056	1.212
0.50	0.235	0.357	0.545	0.782	0.253	0.399	0.630	0.952	0.334	0.518	0.774	1.098	0.375	0.598	0.870	1.182
1.00	0.203	0.312	0.445	0.647	0.244	0.397	0.618	0.934	0.368	0.625	0.969	1.329	0.278	0.455	0.749	1.120
Depth/m	SA-NLNN															
	homogeneous				heterogeneous				two-layered				root water uptake			
	1d	3d	7d	15d	1d	3d	7d	15d	1d	3d	7d	15d	1d	3d	7d	15d
0.05	0.328	0.470	0.686	1.039	0.363	0.524	0.750	1.840	0.327	0.505	0.836	2.210	0.536	0.918	2.150	6.702
0.10	0.249	0.375	0.580	0.957	0.220	0.314	0.477	0.851	0.390	0.569	0.808	1.465	0.322	0.447	0.675	1.480
0.20	0.262	0.366	0.519	0.820	0.292	0.389	0.482	0.648	0.487	0.696	0.945	1.350	0.379	0.546	0.775	1.861
0.50	0.209	0.291	0.414	0.566	0.265	0.337	0.431	0.623	0.327	0.483	0.708	1.018	0.344	0.485	0.687	1.502
1.00	0.245	0.376	0.575	0.807	0.282	0.430	0.640	0.941	0.336	0.530	0.810	1.250	0.297	0.482	0.820	1.748
Depth/m	LSTM 4															
	homogeneous				heterogeneous				two-layered				root water uptake			
	1d	3d	7d	15d	1d	3d	7d	15d	1d	3d	7d	15d	1d	3d	7d	15d
0.05	0.009	1.503	2.791	3.874	0.010	1.520	2.829	3.936	0.015	1.422	2.795	4.106	0.020	1.649	3.187	4.641
0.10	0.007	1.176	2.237	3.184	0.007	1.202	2.282	3.240	0.015	1.479	2.879	4.179	0.014	1.255	2.449	3.584
0.20	0.008	0.786	1.630	2.380	0.010	0.782	1.628	2.384	0.012	0.836	1.735	2.561	0.013	0.801	1.671	2.479
0.50	0.006	0.406	0.942	1.483	0.008	0.373	0.872	1.375	0.008	0.400	0.933	1.476	0.009	0.403	0.939	1.482
1.00	0.008	0.266	0.662	1.116	0.007	0.266	0.664	1.121	0.006	0.258	0.644	1.103	0.006	0.267	0.667	1.136
Depth/m	LSTM 1															
	homogeneous				heterogeneous				two-layered				root water uptake			
	1d	3d	7d	15d	1d	3d	7d	15d	1d	3d	7d	15d	1d	3d	7d	15d
0.05	0.318	0.440	0.590	0.845	0.346	0.484	0.656	0.948	0.264	0.451	0.771	1.313	0.343	0.462	0.600	0.804
0.10	0.135	0.202	0.319	0.528	0.149	0.249	0.408	0.699	0.359	0.542	0.863	1.436	0.274	0.365	0.491	0.707
0.20	0.120	0.174	0.262	0.444	0.138	0.217	0.359	0.669	0.218	0.320	0.545	1.072	0.159	0.238	0.366	0.594
0.50	0.128	0.177	0.274	0.494	0.142	0.207	0.341	0.640	0.179	0.293	0.542	1.090	0.173	0.276	0.443	0.742
1.00	0.214	0.350	0.594	1.075	0.188	0.288	0.465	0.865	0.180	0.293	0.578	1.343	0.242	0.393	0.672	1.203

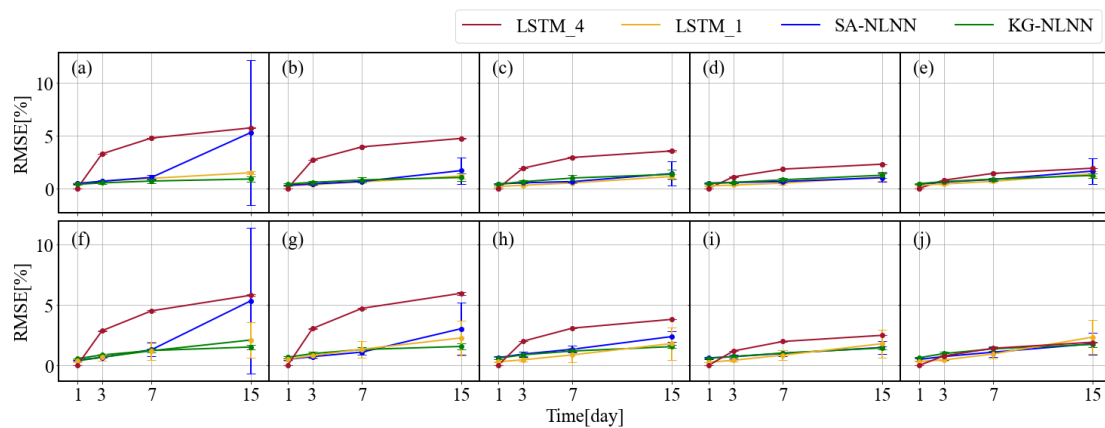


Figure R1. The RMSE results for 1, 3, 7, and 15-day for heterogeneous soil(a-e), and two-layered soil (f-j). The error bar indicates the standard deviations of the RMSE, which are computed via ten training replicates.

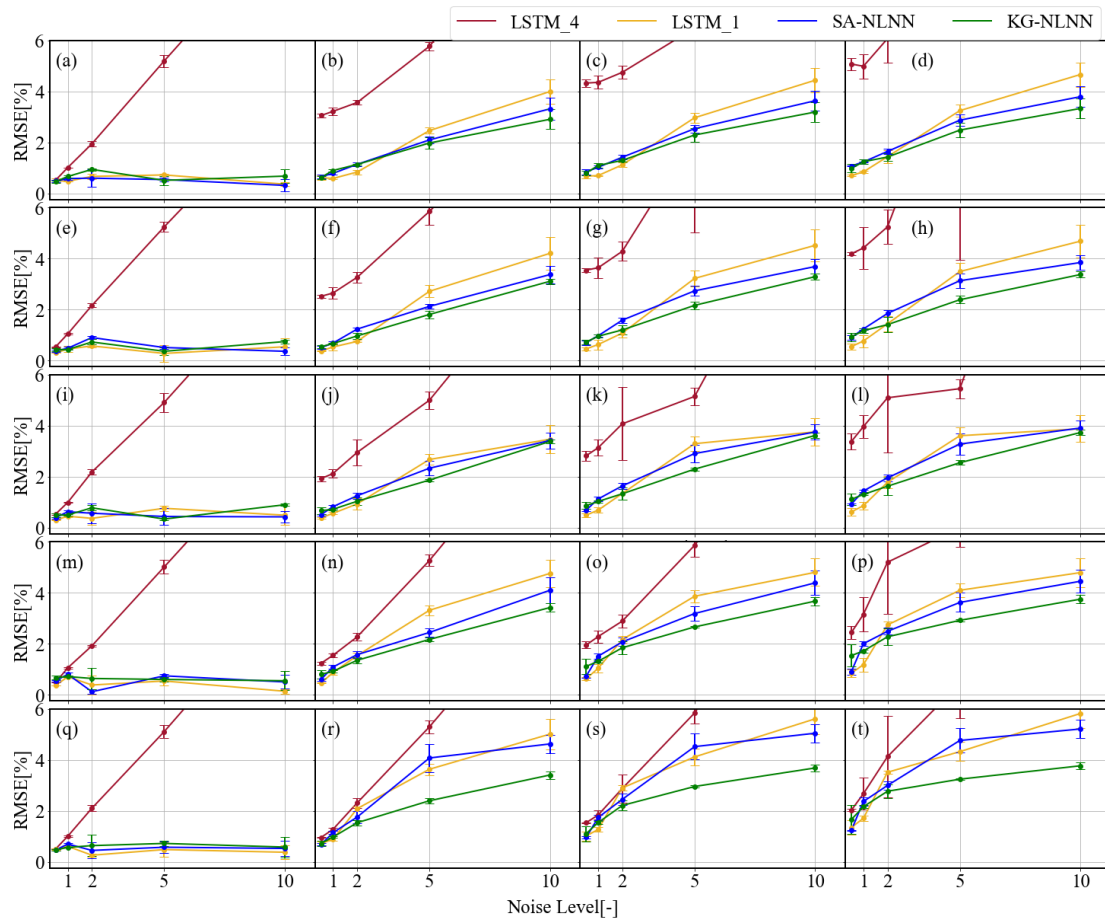


Figure R2. The RMSE results for 1, 3, 7, and 15-day at 0.05m(a-d), 0.10m(e-h), 0.20m(i-l), 0.50m(m-p) and 1.00m(q-t) for homogenous soil under increasing noise levels of the LSTM_4, LSTM_1, SA-NLNN and KG-NLNN models. The error bar indicates the standard deviations of the RMSE, which are computed via ten training replicates.

To improve readability and focus on model interpretability in section 4.1-- given that the extensive statistical data in Table R1 could be overwhelming for synthetic data analysis-- we have moved Table R1 to the appendix. Therefore, Figure R1 now represents the models' comparative accuracy and replaces the original Figure 6. The original Figure 7 has been removed. Furthermore, Figure 14 in the manuscript has been replaced by Figure R2 to provide a more comprehensive evaluation of noise resistance, including a direct comparison with LSTM models.

Interpretability not yet quantitatively tied to physics

Weight maps qualitatively reflect layering, but the link to soil texture/parameters is not quantified (e.g., correlation with (Ksat) contrasts or van Genuchten parameters across cases/sites).

Response2:

Thank you for your comments.

As we mentioned in introduction, our models are not based on any parameterizations and assumptions, instead attempt to learn the soil water dynamics directly from the data. Therefore, it is difficult to yield quantitative descriptions of van Genuchten parameters from the weight maps. Soils

with different van Genuchten parameters (such as K_s) will also exhibit distinct patterns in the weight map. We will provide the following two examples to demonstrate that differences in soil parameters are reflected in the weight map:

Figure R3 (Figure 8 in the manuscript) depicts the weight matrix maps generated by the KG-NLNN model for two-layered soil scenarios. This case is created by interchanging the properties of the upper and lower soil layers, with K_s values set at 0.25 and 10.49, respectively, and the rest of the parameters identical, as recorded in Appendix A. The saturated hydraulic conductivity of the two soil types varies significantly, with distinct characteristics influencing water transport and drainage. With sand below loam, water from the upper loam percolates quickly and deeply, making lower-layer moisture dominated by the upper loam. As shown in Figure R3 (a), the moisture in the lower layer (0.10m, 0.20m, 0.5m, 1.0m) is notably influenced by the moisture at 0.05m. Conversely, with sand above loam, the upper sand rapidly drains water, and the water from the upper sand is absorbed and held by the lower loam. Therefore, soil moisture in the lower layer is mainly affected by the adjacent upper layer, as shown in Figure R3(b). This layered pattern, evident in the weight map, serves as a qualitative indicator of soil texture. Although this cannot provide a quantitative description of the soil hydraulic parameters, it can reflect the difference in hydraulic conductivity between the layers and reveal which layer is more permeable.

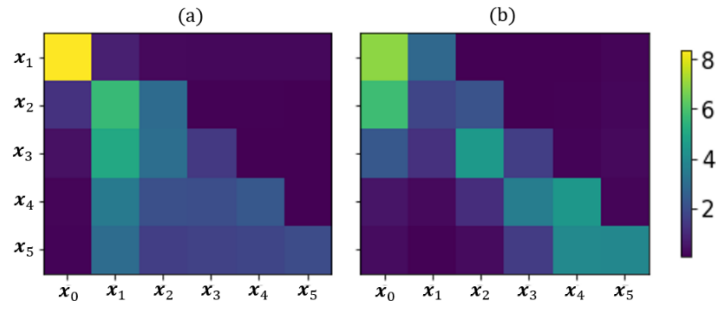


Figure R3. The non-local weight maps in two-layered simulated stratified soil scenarios through KG-NLNN (a) loam above sand (b) sand above loam.

Additionally, we have included a new case in the synthetic data that compares the weight maps of homogeneous soils with different K_s values. The corresponding non-local weight maps derived from KG-NLNN are shown in Figure R4. Differing hydraulic conductivities governs soil water flow velocity, which causes variations in the time required for water to reach various depths. This shapes the formation of the weight maps, leading to the distinctive patterns observed in Figure R4(a) and (b). For instance, loam ($K_s = 0.25$) exhibits slow infiltration, so its moisture content is easily influenced by adjacent layers in Figure R4(a). In contrast, sand ($K_s=10.49$) allows rapid infiltration, resulting in deeper soil moisture being affected directly by meteorological factors. Thus, although our model does not involve any parameterization nor perform a quantitative description of soil hydraulic parameters, it nevertheless provides insights into these hydraulic properties to some extent.

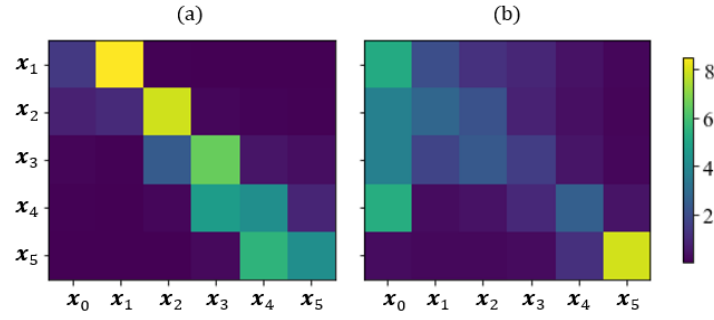


Figure R4. The non-local weight maps in homogeneous simulated soil scenarios through KG-NLNN (a) $K_s = 0.25$ (b) $K_s = 10.49$.

Since our model is not based on any parameterization, our objective is not to quantitatively describe parameters like K_s . Rather, it is to use figures to reflect differences between soils and represent their true states. We have revised the section 4.1 interpretability analysis according to the above.

Minor points

Line 159: ‘ $sm_1 \dots sm_n$ ’ are not explicitly defined.

Response3:

Thank you for comments. Here we have provided a clearer definition of sm_n^t in the manuscript: “Here, sm_n^t denotes the soil moisture at depth n and time t .”

Section 2: How did the authors determine the initial soil moisture for the synthetic and field data cases? Please specify exactly how the first step of multi-day forecasts is initialized in both the synthetic and field experiments

Response4:

Thank you for comments.

Initial Conditions and Data Generation

For the synthetic data generating, we set the initial soil moisture content within the top 1 meter to 0.1 to simplify the conditions. The dataset is then evolved under prescribed meteorological conditions with a free-drainage lower boundary condition. We have revised the sentence here:

“The lower boundary condition is set as free drainage, and the initial moisture content of the soil column is set to a uniform value of 0.10.”

In the case of real-world scenarios, the initial soil moisture value is taken as the 24-hour average measured at the site on the starting date of the selected time period

.

Initialization for Prediction

For model prediction, the models’ initial moisture content input is the truth soil moisture value from the day before the prediction. In the synthetic scenario, this value comes from the physical model’s output; in the real-world scenario, it is taken from field observations. We have added the following description in the manuscript:

“The initial soil moisture content for the prediction is set to the truth from the preceding day.

Specifically, this value is obtained from the physical model’s output for the virtual scenario and from

field observations for the real-world scenario.”