The study conducted a comparison of ten different network structures to assess their predictive abilities and computational costs across various soil textures and depths. The results indicate that Long Short-Term Memory (LSTM), feature attention LSTM (FA-LSTM), and generative adversarial network-based LSTM (GAN-LSTM) are effective in soil moisture forecasting. The study also provides insights into the interpretability of the models and emphasizes the importance of appropriate model design for specific soil moisture prediction tasks. Therefore, this study can serve as a valuable reference for the application of deep learning models in soil water dynamics. Overall, the manuscript is well-organized and easy to follow.

However, there are a few minor issues that the authors should consider. Firstly, it would be beneficial to provide a more detailed description of the representativeness of the sites to avoid potential one-sided conclusions.

Response:

We sincerely thank the reviewer for providing such insightful and detailed comments which have greatly improved the quality of the manuscript. Regarding the reviewer's concern about the data description, we have reorganized Section 2 and added detailed station land cover and meteorology information. Moreover, we have conducted a Pearson correlation analysis for screening input variables.

Response 1:

Table 1 presents the comprehensive details for ten selected sites, sorted from high to low soil permeability. These sites are carefully chosen to illustrate the model's generalization ability, and they encompass ten different soil textures and five distinct land cover types. In addition to site basic meteorology information, Table 2 provides a record of climate data for these selected locations. This data includes minimum, maximum, average, and standard deviation values for air temperature and precipitation. Furthermore, our input data correlation analysis in Figure R1 also demonstrates the variations between the stations.

Table 1. Summary of main characteristics of ten sites.

	Sand	Silt	Clay	Land cover	Period	Lat.	Lon.
Monahans-6-ENE	83	6	11	Shrubland	2010-2022	31.62	102.81
Necedah-5-WNW	83	11	6	Grassland	2009-2022	44.06	-90.17
Falkenberg	73	21	6	Cropland, rained	2003-2020	52.17	14.12
AAMU-jtg	53	22	25	Grassland	2010-2022	34.78	-86.55
Cullman-NAHRC	49	27	24	Mosaic Cropland	2006-2022	34.20	-86.80
Cape-Charles-5-ENE	49	27	24	Herbaceous cover	2011-2022	37.29	-75.93
LittleRiver	47	30	23	Grassland	2005-2020	31.50	-83.55
Spickard	35	41	24	Grassland	2010-2022	40.25	-93.72
Weslaco	34	45	21	Cropland, rained	2017-2021	26.16	-97.96
UpperBethlehem	32	38	30	Herbaceous cover	2008-2010	17.72	-64.80

Table 2. Statistical results of P, and TA at 10 station sites

		Min	Max	Mean	Std	Training set	Validation set	Test set
Monahans-6-ENE	P	0	80.6	0.85	4.60	2010.04.21-	2017.08.25-	2020.02.05-
	TA	-12.78	36.53	19.18	8.86	2017.08.25	2020.02.05	2022.07.19
Necedah-5-WNW	P	0	127.6	2.48	7.23	2009.10.13-	2017.08.27-	2020.04.11-
	TA	-28.87	30.47	7.92	11.69	2017.08.27	2020.04.11	2022.11.26
Falkenberg	P	0	35.34	0.73	1.95	2003.01.17-	2013.07.07-	2017.01.01-
	TA	-18.19	29.45	9.69	7.82	2013.07.07	2017.01.01	2020.06.30
AAMU-jtg	P	0	175.26	2.44	9.42	2010.02.06-	2017.10.07-	2020.04.27-
	TA	-10.83	31.27	16.69	8.24	2017.10.07	2020.04.27	2022.11.18
Cullman-NAHRC	P	0	177.28	2.18	7.73	2006.05.18-	2016.04.19-	2019.08.10-
	TA	-10.07	30.61	16.00	8.28	2016.04.19	2019.08.10	2022.11.30
Cape-Charles-5-	P	0	159.10	2.94	9.19	2011.06.15-	2018.04.13-	2020.07.22-
	TA	-10.47	32.11	15.67	8.53	2018.04.13	2020.07.22	2022.11.01
LittleRiver	P	0	154.68	2.95	9.62	2005.10.18-	2014.04.26-	2017.02.26-
	TA	-4.24	31.99	19.77	7.08	2014.04.26	2017.02.26	2020.01.01
Spickard	P	0	152.91	2.43	8.59	2010.10.08-	2018.01.18-	2020.06.22-
	TA	-22.13	32.31	11.64	11.17	2018.01.18	2020.06.22	2022.11.26
Weslaco	P	0	294.89	1.65	11.66	2017.01.01-	2019.08.07-	2020.06.18-
	TA	-1.41	32.46	23.46	6.07	2019.08.07	2020.06.18	2021.05.01
UpperBethlehem	P	0	156.20	2.78	10.12	2008.09.15-	2009.09.05-	2010.01.01-
	TA	21.64	28.78	25.93	1.46	2009.09.05	2010.01.01	2010.05.01

Additionally, conducting a sensitivity analysis for the input factors would provide further justification for the input screening of the deep learning models.

Response 2:

In the process of screening input factors, we have carefully selected meteorological inputs based on the precipitation and evapotranspiration calculation. Besides, soil temperature data, along with soil moisture data from the previous day are incorporated to represent the soil condition. Figure R1 displays the Pearson correlation analysis results for input factors at the Cape-Charles and UpperBethlem sites. Notably, the correlation coefficients between soil moisture data and the input data vary greatly with both the station and depth. For instance, while the correlation coefficient between longwave radiation (LW) and soil moisture is low at UpperBethlem, it is significant at Cape-Charles, highlighting the influence of site-specific differences. Although utilizing highly correlated factors as inputs appears to be a logical choice, achieving uniformity across different sites and depths can be a complex task. However, this presents a crucial aspect to explore when evaluating and comparing the performance of models for self-learning screening of significant influencing factors. Therefore, we have chosen to include all eight of these data points as inputs. Figure R2 shows the autocorrelation analysis conducted at 5 soil depths. The autocorrelation coefficients for soil water content at different depths decrease with increasing delay days. The most significant change is observed in the surface layer. As a result, we have opted to use a 4-day delay as our input.

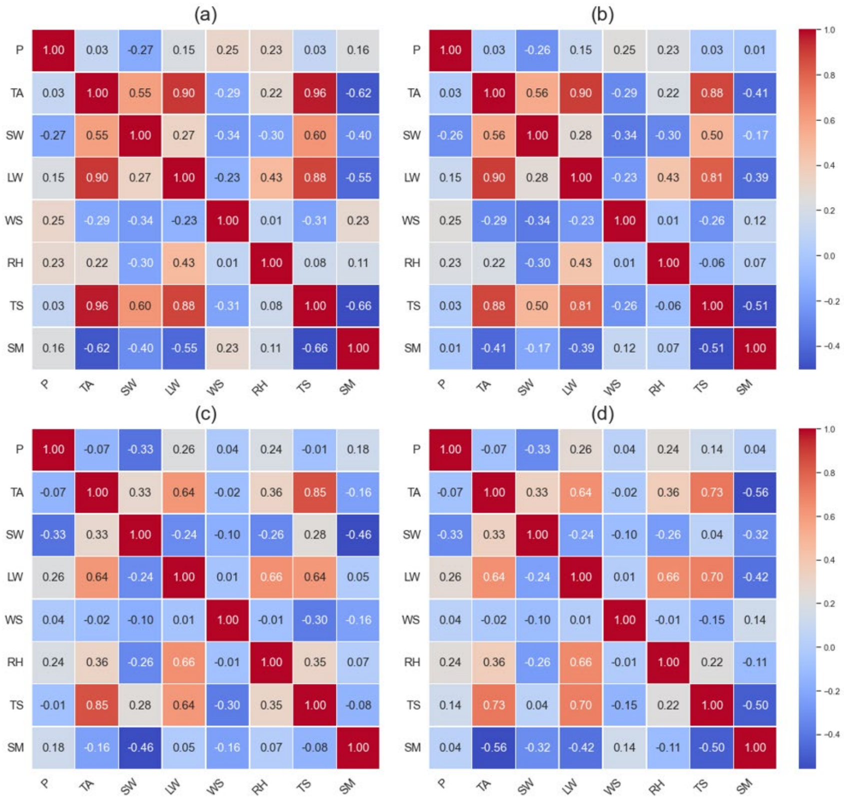

Figure R1. Pearson correlation analysis results among the observed variables of 0.05m and 1.00m at Cape-Charles (a) (b) and UpperBethlem (c) (d) sites.

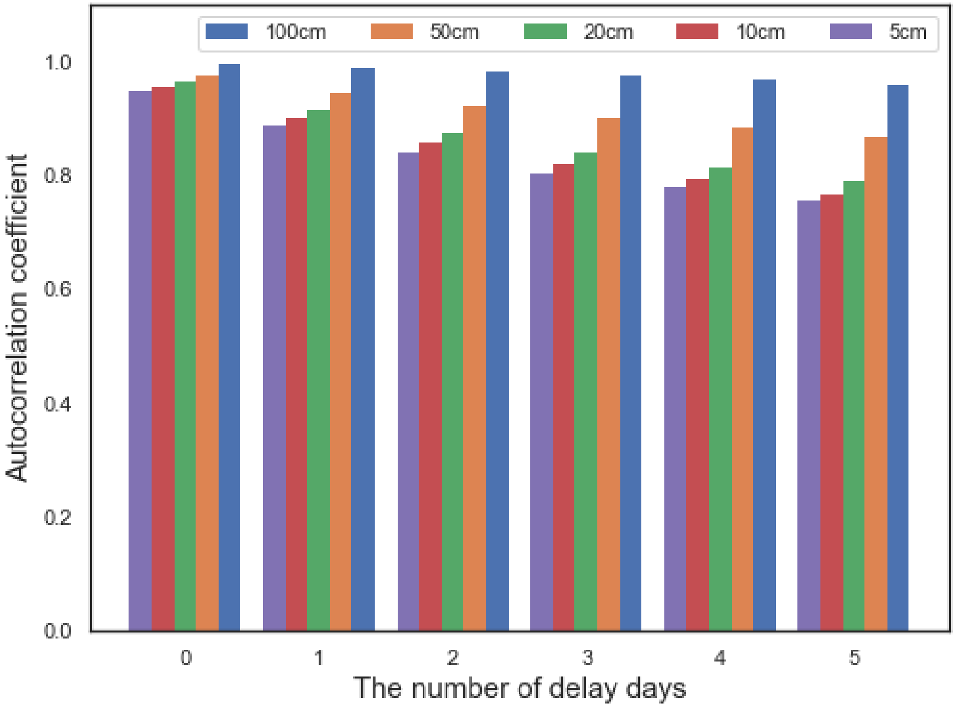

Figure R2. Autocorrelation analysis results of soil water content with different days delay at Cape-Charles