

Local hydrological characteristics shape 1–14-day observation-driven soil moisture predictability

Supplementary Materials

This file includes the following:

Supplementary Methods

Tables S1–S10

Figures S1–S17

Supplementary References

Supplementary Methods

Methods S1. Base learner settings

To leverage the complementary strengths of diverse ML paradigms for SM prediction, we integrated 18 base models spanning four major algorithmic categories: linear, instance-based, tree-based, and neural network approaches.

Linear models: generalized linear model (GLM), Bayesian ridge (BY), and elastic net (EN) were used to represent background linear soil-water redistribution under steady-state matrix flow conditions.

Instance-based model: Support Vector Regression (SVR) was utilized to project inputs into higher-dimensional spaces, capturing highly localized nonlinear structures.

Tree-based models: Decision Tree (DT), Random Forest (RF), Gradient Boosting Regression (GBR), Extreme Gradient Boosting (XGB), and Light Gradient Boosting Machine (LGBM) offered built-in feature ranking, robustness to missing values, and strong performance with threshold-dependent and nonlinear hydrological processes, though they possess limited capacity to extrapolate or encode continuous temporal dependencies.

Neural network models: Backpropagation Neural Network (BPNN), Recurrent Neural Network (RNN), Temporal Convolutional Networks (TCN, Fig. S1), Long Short-Term Memory (LSTM, Fig. S2a), Gated Recurrent Unit (GRU), LSTM with attention, GRU with attention, bidirectional LSTM (BiLSTM, Fig. S2b), and Encoder-Decoder LSTM (EDLSTM). These Neural Network (NN) structures and settings can be found in Table S2. These NNs are specifically designed to capture temporal dynamics, memory effects and non-stationary lagged behaviors in SM evolution. The integration of attention mechanisms further enables the models to identify and weight key meteorological drivers according to their relevance at different lead times. Model settings of the models are provided in Table S3.

To evaluate the optimized stacking ensemble, we compared it with an averaging ensemble and a best-single strategy. The averaging ensemble combined the outputs (Preds) from all 18 base learners through the arithmetic mean. This strategy reduces variance through error cancellation and yields a stable, conservative forecast, but lacks adaptability to changing hydroclimatic conditions. The best-single strategy retained the best-performing individual base learner as the comparator. To ensure a fair comparison, both benchmark strategies followed an identical data exposure schedule. The base models served as pre-trained architectures that were further fine-tuned using the same middle 40% data segment, with their final performance evaluated on the remaining 10% partition. Their performance relative to the stacking feature configurations is presented in Fig. S7.

Methods S2. Richards–VGM diagnostic flux partitioning

Daily gravity-driven and capillary/diffusive flux components were diagnostically estimated from in-situ SM observations using the van Genuchten–Mualem hydraulic functions (Mualem, 1976; van Genuchten, 1980). SoilGrids texture fractions over 0–60 cm were thickness-weighted and classified into USDA texture classes, from which (θ_r), (θ_s), (α), (n), and (K_s) were assigned using the lookup values of Carsel and Parrish

(1988). Shallow and deep SM were approximated as the means of observations at 10–30 and 40–60 cm, respectively.

Effective saturation, matric head, and unsaturated hydraulic conductivity were calculated as

$$S_e = \frac{\theta - \theta_r}{\theta_s - \theta_r}, \quad m = 1 - \frac{1}{n} \# (S1)$$

$$|\psi| = \frac{1}{\alpha} (S_e^{-1/m} - 1)^{1/n}, \quad K(\theta) = K_s S_e^{1/2} [1 - (1 - S_e^{1/m})^m]^2 \# (S2)$$

For numerical stability, (S_e) was restricted to $[10^{-5}, 0.999999]$.

Interface conductivity was estimated as the geometric mean of the shallow- and deep-layer conductivities. Using a 30 cm vertical separation, the gravity-driven and capillary/diffusive components were defined as

$$K_{int} = \sqrt{K_{shallow} K_{deep}}, q_g = K_{int}, q_c = K_{int} \frac{|\psi_{deep} - \psi_{shallow}|}{30} \# (S3)$$

Rainfall days were defined by daily precipitation (≥ 1) mm; the remaining days were classified as non-rainfall days.

Daily estimates outside the assigned (θ_r, θ_s) range were excluded. Median (q_g) and (q_c) values were then calculated for each site and rainfall condition and compared between FD and SD regimes using paired Wilcoxon signed-rank tests on log-transformed fluxes.

Methods S3. Entropy-based validation of predictability regimes

To independently validate whether the predictability regimes identified from $\text{NRMSE}_{\text{std}}$ trajectories reflect intrinsic differences in local SM dynamics, we utilized SampEn to analyze the in-situ time series. First, SampEn was calculated from daily SM observations for each site and soil depth (10, 30, and 60 cm) using a standard configuration (embedding dimension=2, tolerance= $0.2 \times \sigma$ of SM time series) (Chou, 2012). Higher SampEn values indicate lower temporal regularity and greater SM dynamic complexity. Second, to ensure the robustness of our findings, we evaluated SampEn under alternative configurations, including: 1) varying tolerance levels (tolerance=0.15, 0.20, and $0.25 \times \sigma$); 2) using first-difference SM series to remove long-term trends and seasonality; and 3) calculating multiple scale SampEn at 1-, 3-, and 7-day scales (Richman and Moorman, 2000; Costa et al., 2002, 2005). Higher SampEn values indicate greater temporal irregularity and lower predictability of SM dynamics. Additionally, the Mann–Whitney U test was applied to determine whether SampEn values differed significantly between the two identified regimes, while the ROC-AUC score was used to assess SampEn's discriminative power in distinguishing these regimes. Furthermore, we employed Spearman correlation analysis to examine the consistency between the entropy metrics and the $\text{NRMSE}_{\text{std}}$ growth rate across sites. To account for potential confounding factors, a partial Spearman correlation was subsequently performed, controlling for the log-transformed record length of each

station.

Supplementary Tables

Table S1. In situ soil moisture observations in this study.

Site ID	Latitude	Longitude	Region	Altitude	Site ID	Latitude	Longitude	Region	Altitude	Site ID	Latitude	Longitude	Region	Altitude
1	25.0589	113.7631	North	105	26	21.65	110.6167	West	18	51	21.680956	110.513025	West	23
2	25.0806	114.2561	North	138	27	21.6108	111.5753	West	6	52	22.132797	110.664683	West	62
3	24.8	112.3667	North	105	28	20.3369	110.1961	West	48	53	22.539186	111.149717	West	587
4	24.95	114.0667	North	112	29	23.1347	113.0458	PRD	15	54	23.513372	111.813597	PRD	64
5	24.3667	114.4831	North	222	30	25.1333	114.3167	North	135	55	24.276131	116.256708	North	106
6	24.2667	116.1	North	97	31	22.6789	113.3989	PRD	5	56	24.432636	116.08025	North	160
7	23.3975	111.5081	PRD	41	32	23.394181	113.437636	PRD	24	57	24.596481	116.390572	North	314
8	23.1347	113.0458	PRD	15	33	23.303486	113.696256	PRD	15	58	24.16649007	116.7982273	North	762
9	23.6333	112.4333	PRD	40	34	23.601683	113.350372	PRD	44	59	23.465878	115.502281	North	688
10	23.0333	112.45	PRD	19	35	24.163542	114.287244	North	184	60	24.493617	115.911214	North	316
11	23.55	113.5833	PRD	43	36	23.209811	113.070625	PRD	6	61	23.004422	115.37125	East	3
12	23.05	113.1333	PRD	12	37	22.51082058	113.099598	PRD	5	62	23.648847	115.217356	North	247
13	22.9667	113.7333	PRD	26	38	22.043208	112.944083	PRD	9	63	23.504603	114.701511	North	25
14	23.78	114.2489	PRD	81	39	22.395008	112.6747	PRD	2	64	24.14317735	115.312264	North	112
15	23.9333	115.7667	North	104	40	22.514175	112.789989	PRD	1	65	23.05140571	113.5818023	PRD	6
16	23.6333	115.1833	East	139	41	22.379347	112.443697	PRD	10	66	22.63372423	113.5427528	PRD	0
17	23.6833	116.9833	East	10	42	21.18004178	110.3093357	West	86	67	22.359583	113.467697	PRD	9
18	23.3	116.1333	East	21	43	20.4779362	110.3402403	West	131	68	22.702917	113.353222	PRD	0
19	23.2667	116.5833	East	45	44	21.66480333	110.1756372	West	22	69	22.464194	113.546433	PRD	0
20	22.35	110.9333	West	86	45	21.575317	109.894135	West	24	70	22.597111	113.509533	PRD	6
21	22.7	111.6	West	61	46	20.827989	110.170389	West	9	71	23.3524783	116.2741501	East	22

22	22.2 5	112. 7833	PRD	8	47	21.75 93758 1	110.796 4773	West	22	72	23.057 38607	112.195 7878	West	38
23	22.9 4	113. 4453	PRD	5	48	21.56 3153	111.339 708	West	12	73	22.653 367	111.622 708	West	63
24	22.5 333	113. 35	PRD	3	49	21.85 0981	110.710 189	West	29					
25	22.9 5	115. 65	East	45	50	21.89 2392	110.777 522	West	25					

Table S2. Description of input variables for base-model training.

variable type	Variable Name	Description & Units	Data Source
Time Features	seq	Site-specific daily sequence index (1–n; dimensionless)	Observation timestamps
Time Features	doy	Day of year (dimensionless)	Observation timestamps
Soil Moisture	SMi/SMs	10-100cm, soil moisture at 8 layers (%)	In situ Observations
Meteorological variables	tp	Total precipitation (m day ⁻¹)	ERA5-Land
Meteorological variables	r	Relative humidity (%)	ERA5-Land
Meteorological variables	t2m	2m temperature (K)	ERA5-Land
Meteorological variables	sp	Surface pressure (Pa)	ERA5-Land
Meteorological variables	d2m	2m dewpoint temperature (K)	ERA5-Land
Meteorological variables	u10, v10	10m wind components (m s ⁻¹)	ERA5-Land
Meteorological variables	e	Total evaporation (m)	ERA5-Land
Meteorological variables	pev	Potential evaporation (m)	ERA5-Land
Meteorological variables	ssr	Surface net solar radiation (J m ⁻²)	ERA5-Land
Meteorological variables	slhf	Surface latent heat flux (J m ⁻²)	ERA5-Land
Meteorological variables	ro	Total runoff (mm day ⁻¹)	ERA5-Land
Meteorological variables	fal	Forecast albedo (unitless)	ERA5-Land

Table S3. Neural network model structures and settings.

Model short name	Layer-by-Layer Architecture	Key Parameters	Activation
BPNN	Dense -> Dense -> Dense	Units: 64, 32, 1	ReLU, Sigmoid
RNN	RNN -> RNN -> Dense	Units: 100, 50; Dropout: 0.1	ReLU
LSTM	LSTM -> LSTM -> Dense	Units: 64, 32	ReLU
BiLSTM	Bi-LSTM -> Bi-LSTM -> Dense	Units: 64, 32	ReLU
GRU	GRU -> Dense -> Dense	Units: 64, 32; Dropout: 0.2	ReLU
LSTMatt	LSTM -> Attention -> Dense	Units: 128, 32; Dropout: 0.1	ReLU
GRUatt	GRU -> Attention -> Dense	Units: 128, 32; Dropout: 0.1	ReLU
EDLSTM	LSTM (Enc) -> Repeat -> LSTM (Dec) -> Dense	Units: 200, 200, 100	ReLU
TCN	ResBlock *3 -> Flatten -> Dense	Filters: 32, 32, 16; Dilation: 1, 2, 4	ReLU

Table S4. Feature configuration strategies within the stacking meta-model.

Feature strategy	Predictions of base-models (Preds)	SM at the target depth (SMi)	SM at other depths (SMs)	Meteorological variables (MV)	Feature selection (FS)*
S1	√				
S2	√	√			
S3	√	√		√	
S4	√	√	√		
S5	√	√	√	√	√
S6	√	√	√	√	

*The feature selection for S5 was based on the mean SHAP values to remove the correlated and unimportant predictors (Huang et al., 2023).

Table S5. Cross-validated support for the selected Pmax-TWS factor pair.

Analysis	Predictor set/rule	Classifier/test	n	BACC	AUC	Fisher p	Validation scheme	Interpretation
Two-factor CV screening	Pmax + TWS	Balanced logistic regression	62	0.744	0.778		Stratified K-fold CV	Cross-validated separation of the selected rainfall-soil pair.
Two-factor CV screening	Pmax + TWS	Shallow decision tree	62	0.696	0.733		Stratified K-fold CV	Consistent but slightly weaker separation under a threshold-based classifier.
Threshold-domain support	Pmax $\geq$ 27.65 mm and TWS $\leq$ 0.70	Constrained quadrant + Fisher test	62	0.776		1.35e-05	73.5% inside vs 17.9% outside	Statistically supported high-risk rainfall-soil domain.

Table S6. Regime-dependent grouped SHAP contribution proportions in the stacking meta-model, averaged across sites and lead times. Positive differences indicate larger contribution in the FD regime.

Depth	Feature group	SD	FD	FD - SD	Main contributors
10 cm	MV	0.295	0.254	-0.040	seq;fal;doy;sp;Other
10 cm	Preds	0.506	0.586	+0.079	BY;EN;LGBM;GBR;Other
10 cm	SMi+SMs	0.168	0.160	-0.008	10cm;60cm;80cm;20cm;Other
30 cm	MV	0.276	0.205	-0.071	seq;fal;sp;t2m;Other
30 cm	Preds	0.514	0.617	+0.103	EN;BY;LGBM;RF;Other
30 cm	SMi+SMs	0.179	0.178	-0.001	30cm;10cm;20cm;60cm;Other
60 cm	MV	0.326	0.238	-0.089	fal;seq;doy;ro;Other
60 cm	Preds	0.390	0.500	+0.111	BY;EN;LGBM;RF;Other
60 cm	SMi+SMs	0.253	0.239	-0.013	60cm;10cm;80cm;40cm;Other

Table S7. Single-factor, climatic-moment, joint-factor, and geographic-background diagnostics for distinguishing FD and SD regimes. AUC values are direction-free; partial Spearman rows control for the listed variables.

Category	Descriptor or rule	Type	SD median	FD median	Direction	Separability statistic	Statistic value	P value
Climate	Pmax	Single factor	37.63	55.41	higher in FD	AUC	0.710	0.0045
Climate	t2m_min	Single factor	279.9	277.0	lower in FD	AUC	0.693	0.0094
Climate	tp_std	Single factor	4.362	5.064	higher in FD	AUC	0.650	0.0432
Climate	tp_mean	Single factor	2.349	2.606	higher in FD	AUC	0.644	0.0528
Climate	e_max	Single factor	-0.0004	-0.0003	higher in FD	AUC	0.643	0.0536
Climate	pev_max	Single factor	-0.0005	-0.0004	higher in FD	AUC	0.618	0.1114
Climate	t2m_max	Single factor	304.5	304.3	lower in FD	AUC	0.617	0.1146
Climate	e_min	Single factor	-0.0049	-0.0049	lower in FD	AUC	0.589	0.2339
Climate	t2m_std	Single factor	5.609	5.823	higher in FD	AUC	0.579	0.2876
Climate	pev_min	Single factor	-0.0088	-0.0093	lower in FD	AUC	0.578	0.2972
Climate	e_mean	Single factor	-0.0022	-0.0022	higher in FD	AUC	0.553	0.4769
Climate	e_std	Single factor	0.0011	0.0010	lower in FD	AUC	0.546	0.5401
Climate	t2m_mean	Single factor	294.8	294.8	lower in FD	AUC	0.535	0.6370
Climate	tp_min	Single factor	0.0000	0.0000	lower in FD	AUC	0.531	0.1746
Climate	pev_mean	Single factor	-0.0041	-0.0040	higher in FD	AUC	0.531	0.6778
Climate	pev_std	Single factor	0.0015	0.0015	higher in FD	AUC	0.514	0.8602
Soil	TWS	Single factor	0.9625	0.0000	lower in FD	AUC	0.724	9.98e-04
Soil	Texture mutation	Single factor	26.96	41.11	higher in FD	AUC	0.697	0.0073
Soil	Mean silt	Single factor	35.87	33.93	lower in FD	AUC	0.685	0.0119
Soil	Field capacity proxy	Single factor	35.91	33.62	lower in FD	AUC	0.583	0.2625

Soil	BDOD std.	Single factor	8.232	8.617	higher in FD	AUC	0.560	0.417 2
Soil	Clay gradient	Single factor	47.00	53.00	higher in FD	AUC	0.527	0.715 6
Geograph y	Altitude Pmax	Partial Spearman			associatio n with FD	rho	0.344	0.006 2
Geograph y	Altitude Pmax + TWS	Partial Spearman			associatio n with FD	rho	0.191	0.137 8
Geograph y	Altitude	Single factor	15.00	44.00	higher in FD	AUC	0.710	0.004 3
Geograph y	Longitud e	Single factor	113.4	113.0	lower in FD	AUC	0.547	0.527 1
Geograph y	Latitude	Single factor	22.85	23.13	higher in FD	AUC	0.541	0.582 4
Data	Length	Single factor	572.5	2529.0	higher in FD	AUC	0.723	0.002 4
Combined	z(Pmax) + z(1 - TWS)	Environmenta l risk score			higher FD risk	AUC	0.805	3.79e- 05
Combined	Pmax >= 27.65 mm and TWS <= 0.70	Rainfall-soil risk domain			higher FD risk	balanced accuracy	0.776	1.35e- 05

Table S8. Robustness of two-factor classification results under random and stratified K-fold validation. AUC values are direction-free. LR = balanced logistic regression; DT = shallow decision tree.

Factor combination	n	Random K-fold AUC (LR)	Stratified K-fold AUC (LR)	Stratified K-fold BACC (LR)	Stratified K-fold AUC (DT)	Stratified K-fold BACC (DT)
TWS + Pmax	6 2	0.779	0.778	0.744	0.733	0.696
TWS + Pstd	6 2	0.786	0.796	0.744	0.792	0.743
TWS + Pmean	6 2	0.780	0.785	0.728	0.758	0.711
TWS + Altitude	6 3	0.690	0.697	0.716	0.683	0.715
Deep SDM + Pmax	6 2	0.755	0.759	0.644	0.719	0.715
Deep SDM + Pstd	6 2	0.741	0.747	0.676	0.764	0.727
Texture mutation + Pmax	6 2	0.748	0.730	0.661	0.713	0.678
Texture mutation + Pstd	6 2	0.669	0.673	0.661	0.655	0.642
Mean silt + Pmax	6 2	0.710	0.733	0.710	0.701	0.661
Mean silt + Pstd	6 2	0.676	0.703	0.708	0.791	0.757
Field capacity + Pmax	6 2	0.723	0.718	0.644	0.672	0.630
BDOD std. + Pmax	6 2	0.700	0.682	0.645	0.579	0.616
Clay gradient + Pmax	6 2	0.695	0.651	0.628	0.599	0.582
Pmax + Altitude	6 2	0.733	0.724	0.724	0.617	0.613

Table S9. Record-length sensitivity checks for the Pmax-TWS rainfall-soil risk domain.

Check	Comparison	Statistic	Result	Interpretation
Single-factor association	Length by regime	Mann-Whitney U; AUC	SD median = 572.5 days; FD median = 2529.0 days; AUC = 0.723; $p = 0.0024$	Record length differs between regimes and therefore requires explicit sensitivity checks.
Repeated-CV comparison	Rainfall-soil model	5-fold x 100 repeats	Balanced logistic regression: BACC = 0.717, AUC = 0.785; shallow decision tree: BACC = 0.671, AUC = 0.716	The Pmax-TWS model provides reproducible regime separation.
Repeated-CV comparison	Rainfall-soil + log(length)	5-fold x 100 repeats	Balanced logistic regression: BACC = 0.698, AUC = 0.784; shallow decision tree: BACC = 0.668, AUC = 0.673	Adding length did not improve cross-validated separation relative to the rainfall-soil model.
Repeated-CV comparison	Length only	5-fold x 100 repeats	Balanced logistic regression: BACC = 0.676, AUC = 0.716; shallow decision tree: BACC = 0.645, AUC = 0.645	Length alone was weaker than the rainfall-soil attribution model.
Length-stratified validation	Shorter records	Risk-domain enrichment	FD rate = 69.2% inside vs 5.6% outside; Fisher $p = 2.97e-04$	The same risk-domain direction is retained within shorter-record sites.
Length-stratified validation	Longer records	Risk-domain enrichment	FD rate = 76.2% inside vs 40.0% outside; Fisher $p = 0.106$	The direction is retained within longer-record sites, although the within-stratum contrast is weaker.

Table S10. Partial Spearman correlation between FDI and $\Delta\text{NRMSE}_{\text{std}}$ under different controlled variables

Controlled variables	n	Partial Spearman r	p value
Record length	62	0.387	0.0021
Altitude	62	0.315	0.0134
Record length + Altitude	62	0.253	0.0509

Supplementary Figures

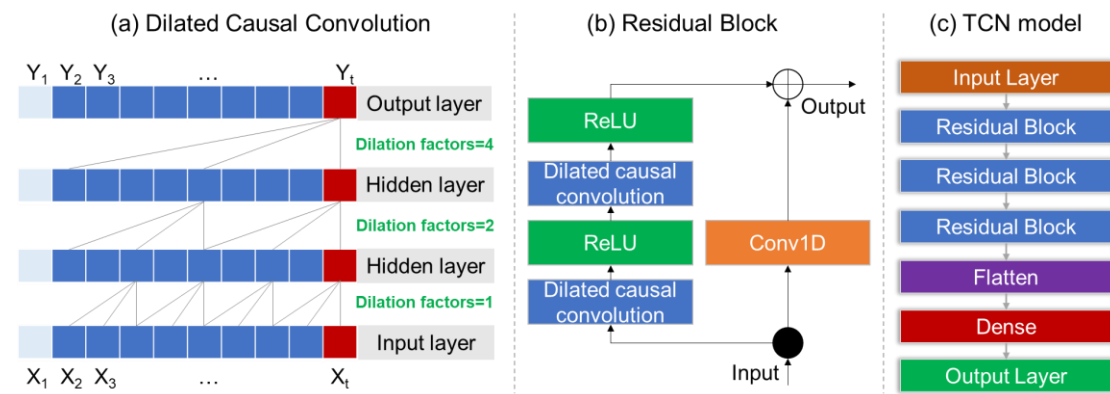


Fig. S1 Temporal convolutional network structure.

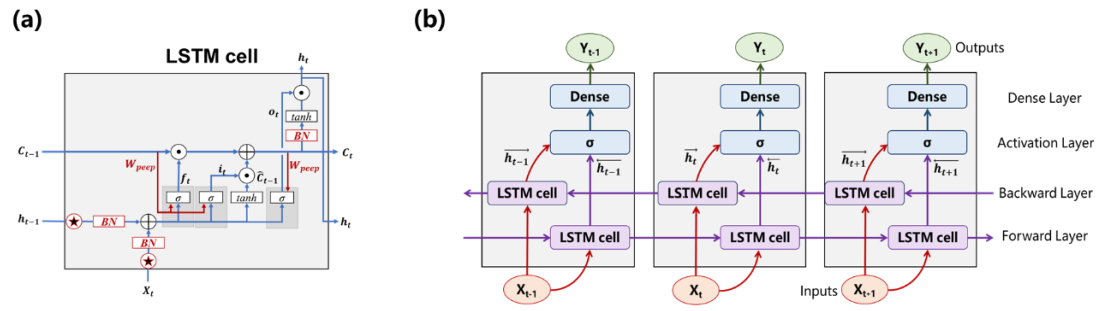


Fig. S2 Long short-term memory (LSTM, a) and bidirectional LSTM (BiLSTM, b) structures.

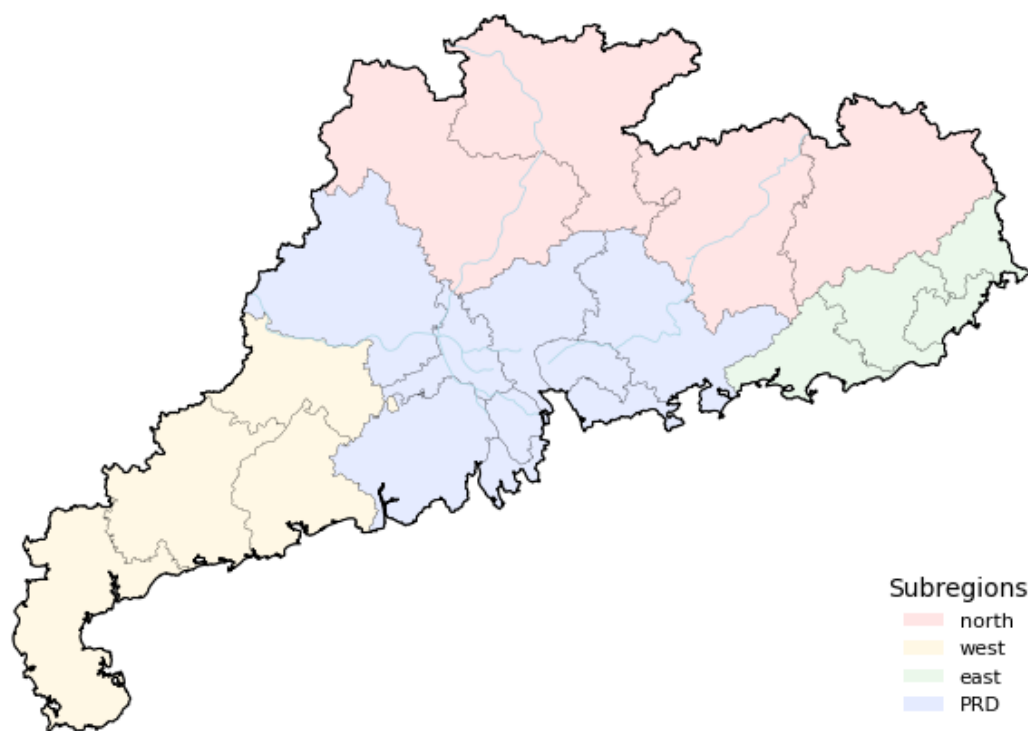


Fig. S3 Sub-regions in Guangdong Province.

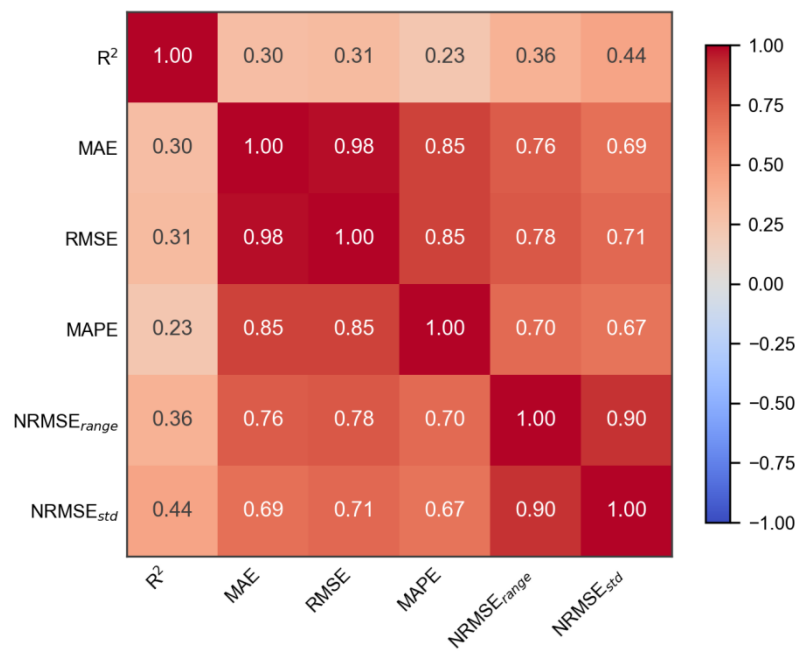


Fig. S4 Performance metric correlation heatmap for stacking forecasts.

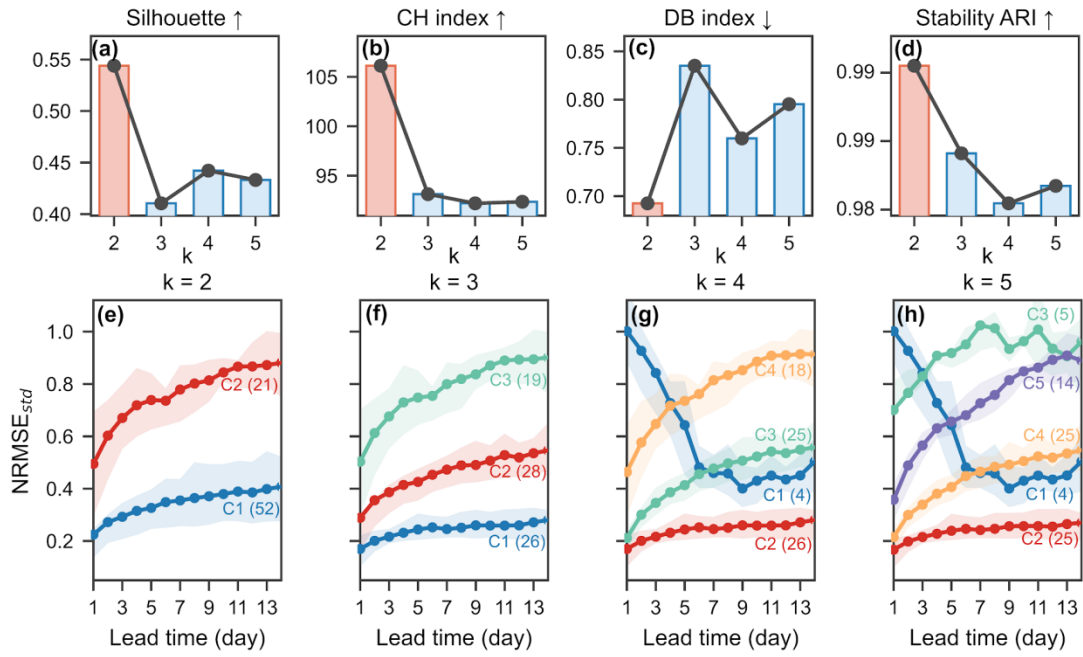


Fig. S5 Selection of optimal cluster number for K-means regime classification. (a-d) Silhouette, Calinski-Harabasz, Davies-Bouldin, and Stability-ARI indices for $k = 2-5$ (red bars: selected $k = 2$). (e-h) Mean NRMSEstd trajectories per cluster at $k = 2-5$.

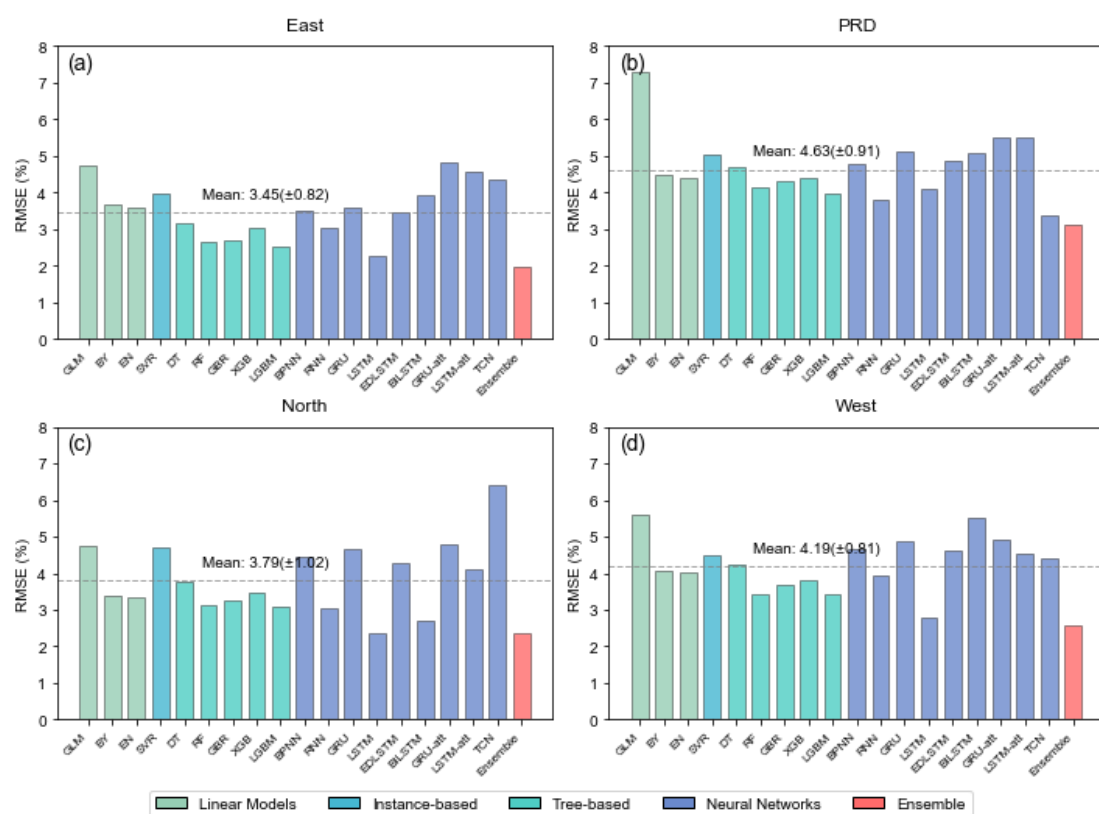


Fig. S6 Performance of base models aggregated by geographic region. Panels (a-d) display mean RMSE (%) for forecasts in East, PRD, North, and West Guangdong, respectively.

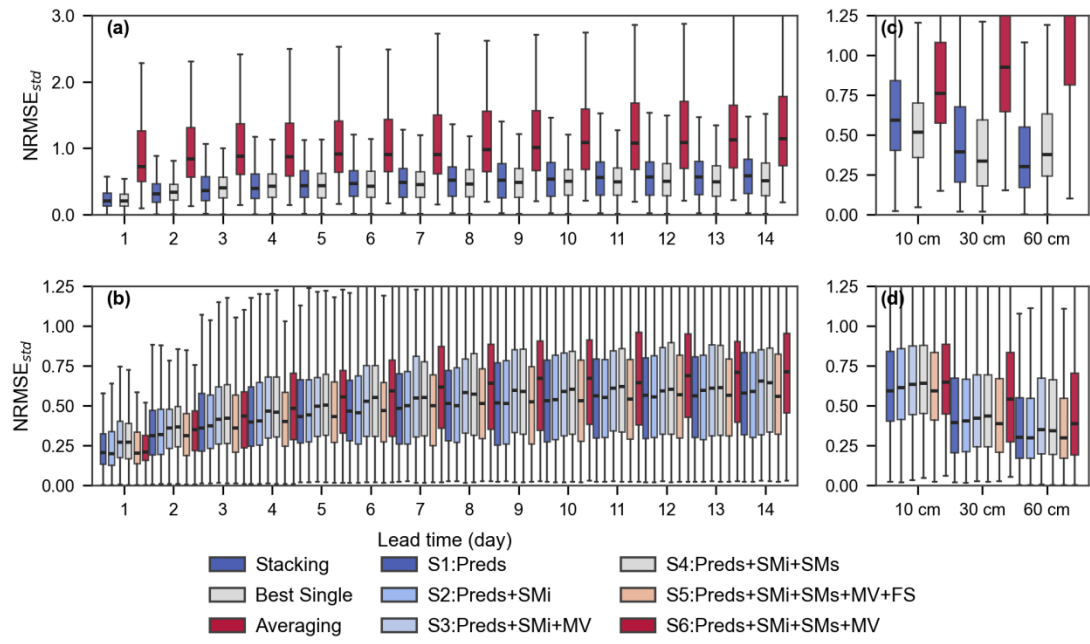


Fig. S7 Ensemble strategy and S1-S6 feature configuration analysis. Panels compare initial ensemble strategies and stacking feature configurations across lead times and soil depths; S6 denotes the final enriched-input stacking strategy.

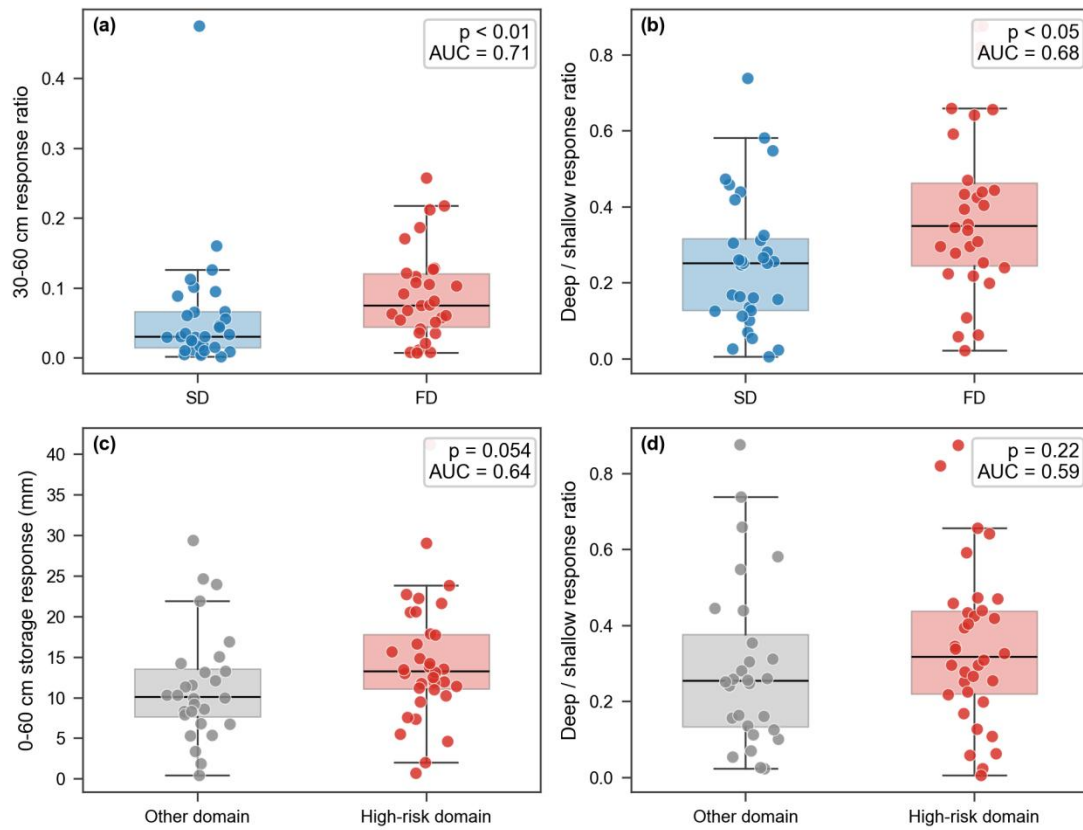


Fig. S8 Event-based soil-water response diagnostics. (a-b) Deep-layer response ratio and vertical redistribution ratio between FD and SD regimes. (c-d) Same metrics between the rainfall-soil high-risk domain and other sites.

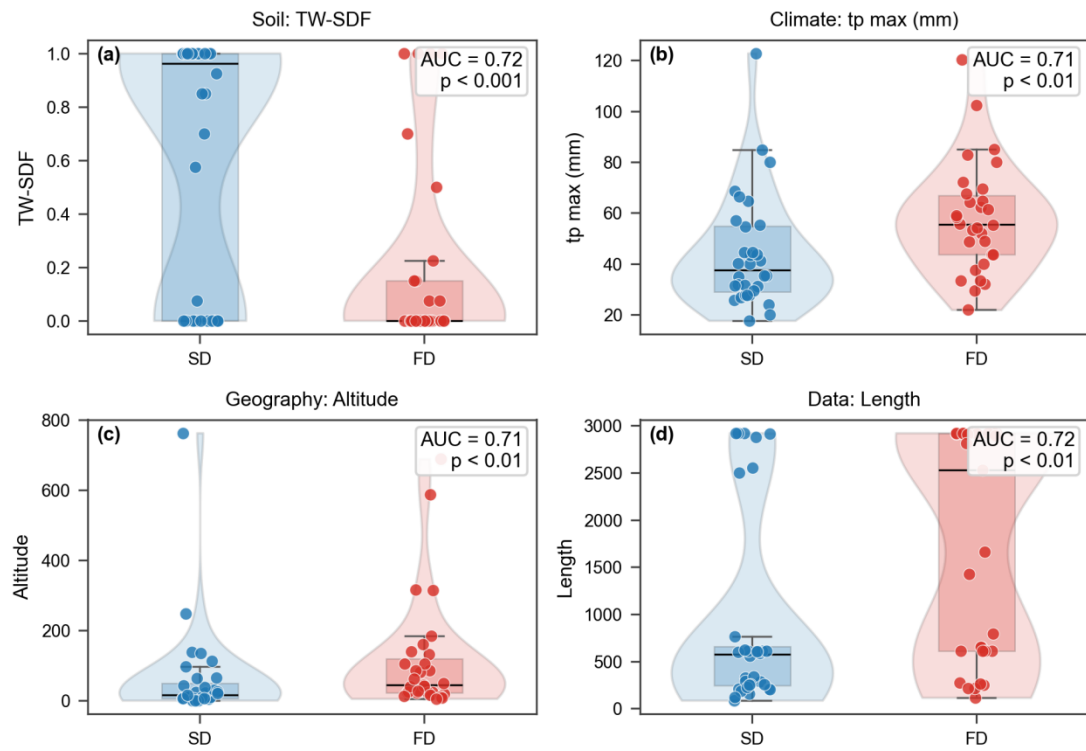


Fig. S9 Distributions of the top soil, climate, geography, and data-length descriptors between FD and SD regimes.

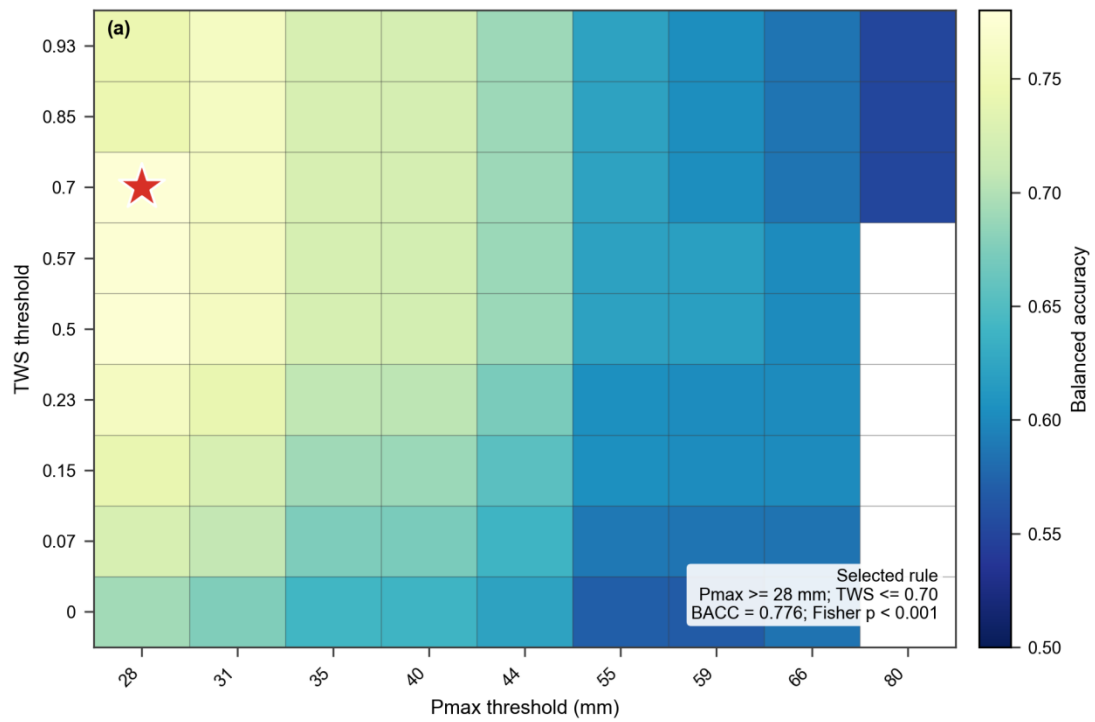


Fig. S10 BACC of the FD regime domain across candidate Pmax (x-axis) and TWS (y-axis) thresholds. The red star marks the selected combination (Pmax ≥ 28 mm, TWS ≤ 0.70 ; BACC = 0.776). White cells had < 5 sites inside or outside the domain.

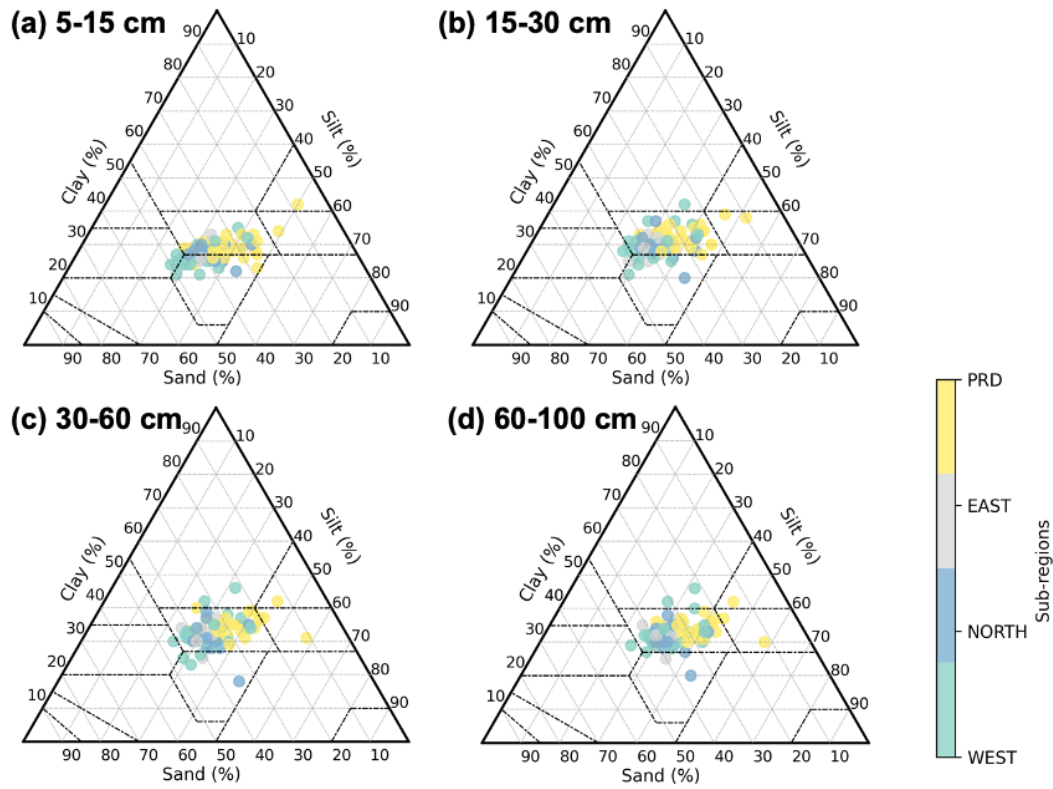


Fig. S11 Soil texture ternary distributions across depth intervals and geographic regions. (a-d) Sand-silt-clay composition at 5-15, 15-30, 30-60, and 60-100 cm, colored by region (East, PRD, North, West). Data: SoilGrids 250 m.

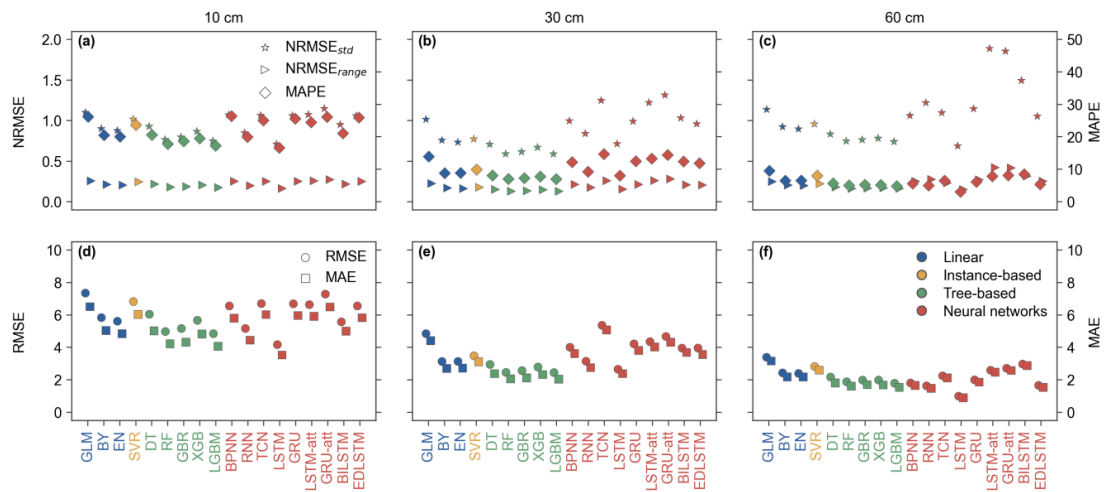


Fig. S12 Performance of the 18 base models across depths, averaged over lead times. Each depth panel shows normalized metrics (NRMSE_{std}, NRMSE_{range}, and MAPE) in the upper axes and absolute errors (RMSE and MAE) in the lower axes. Colors denote model category: linear (blue), instance-based (orange), tree-based (green), neural network (red).

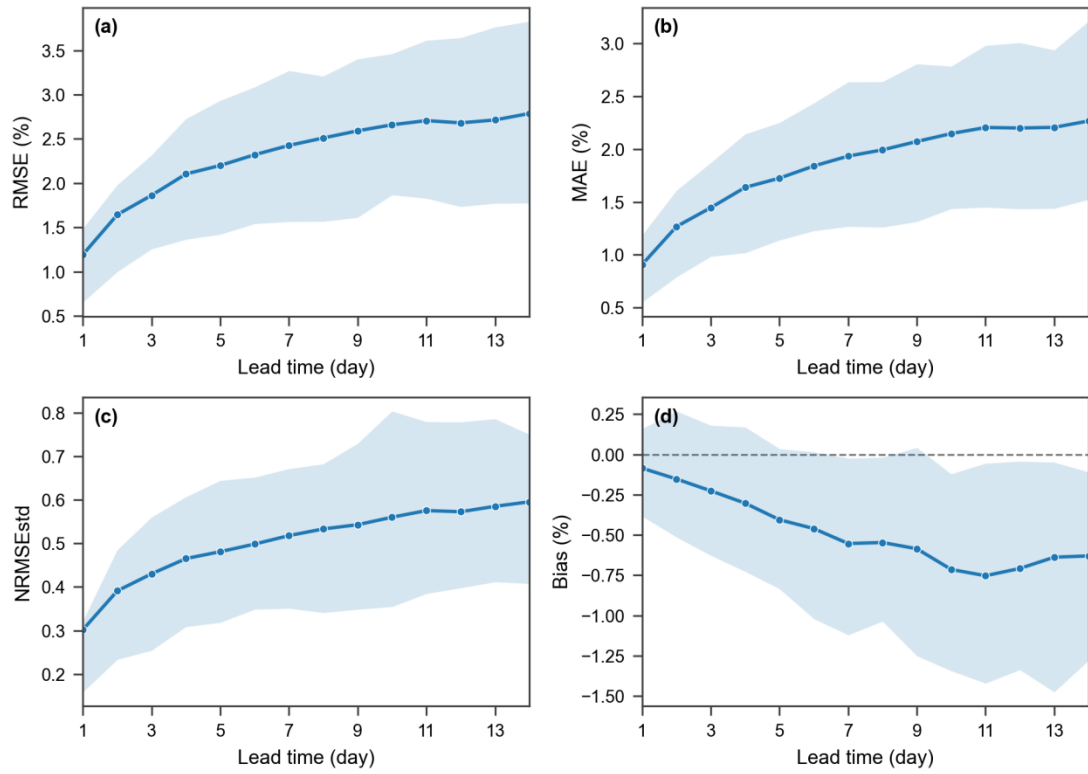


Fig. S13 Lead-time evolution of stacking-ensemble forecast performance across all sites and depths. (a) RMSE, (b) MAE, (c) NRMSEstd, (d) Bias (dashed line: zero).

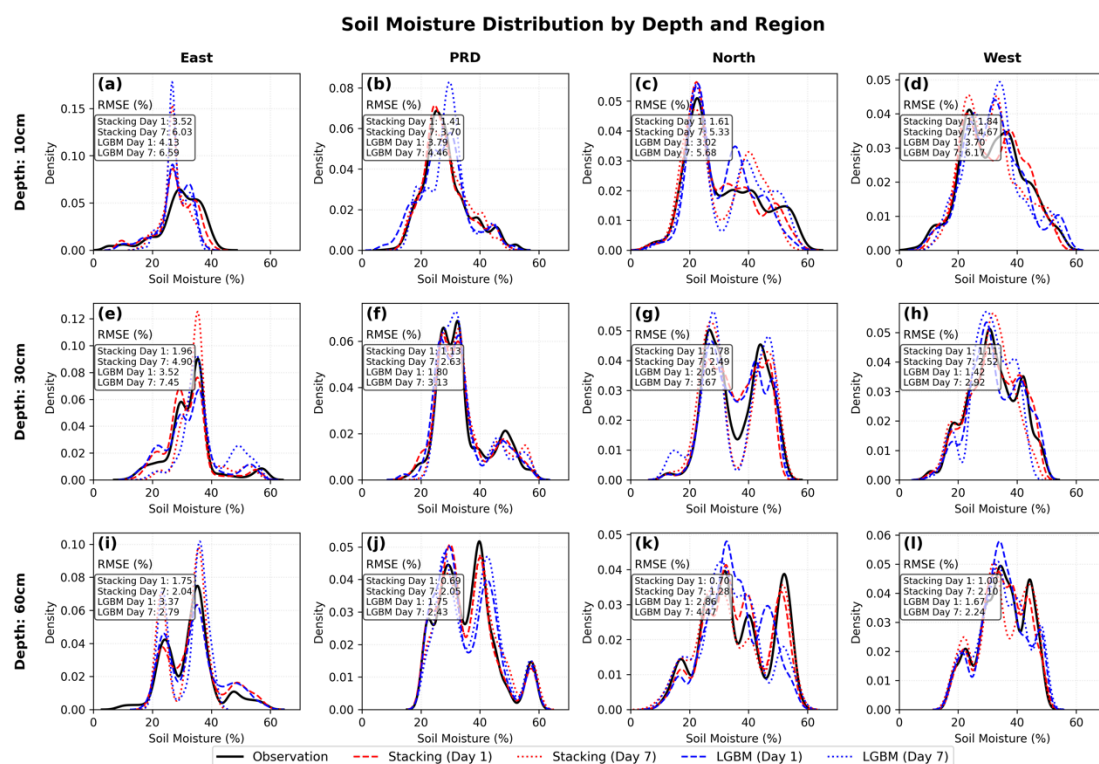


Fig. S14 Regional/depth SM distribution comparisons. Observed SM distributions are compared with S6 stacking and LGBM forecasts at Day 1 and Day 7 across regions and target depths.

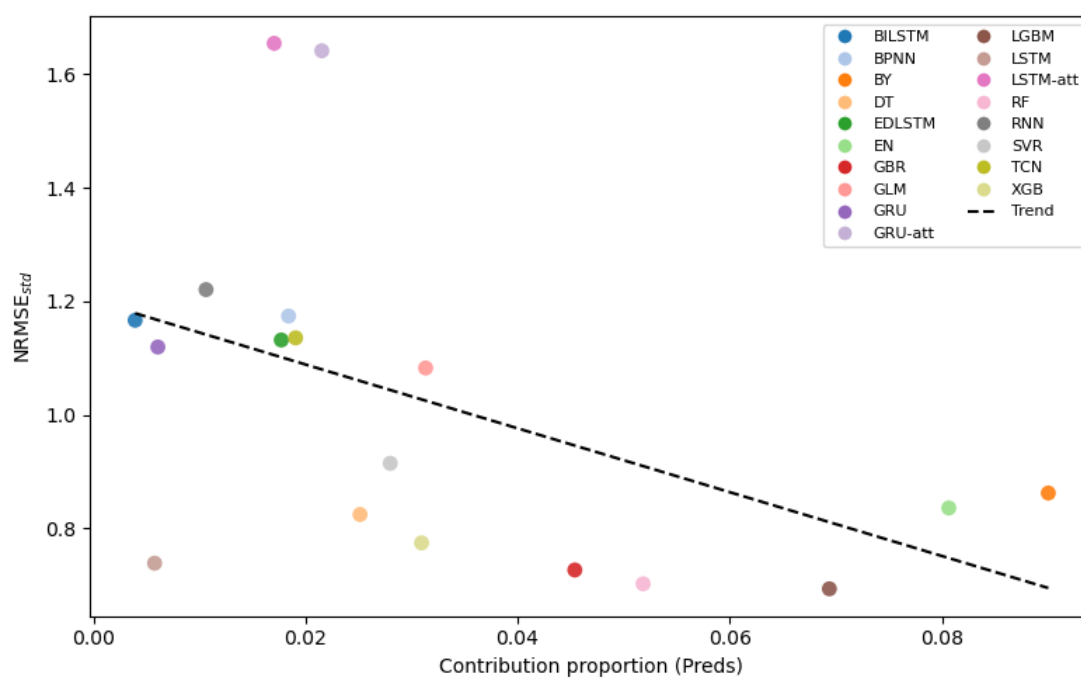


Fig. S15 Relationship between NRMSEstd and contribution proportion estimated by normalized SHAP values in the stacking meta-model. Values are averaged across lead times and SM layers.

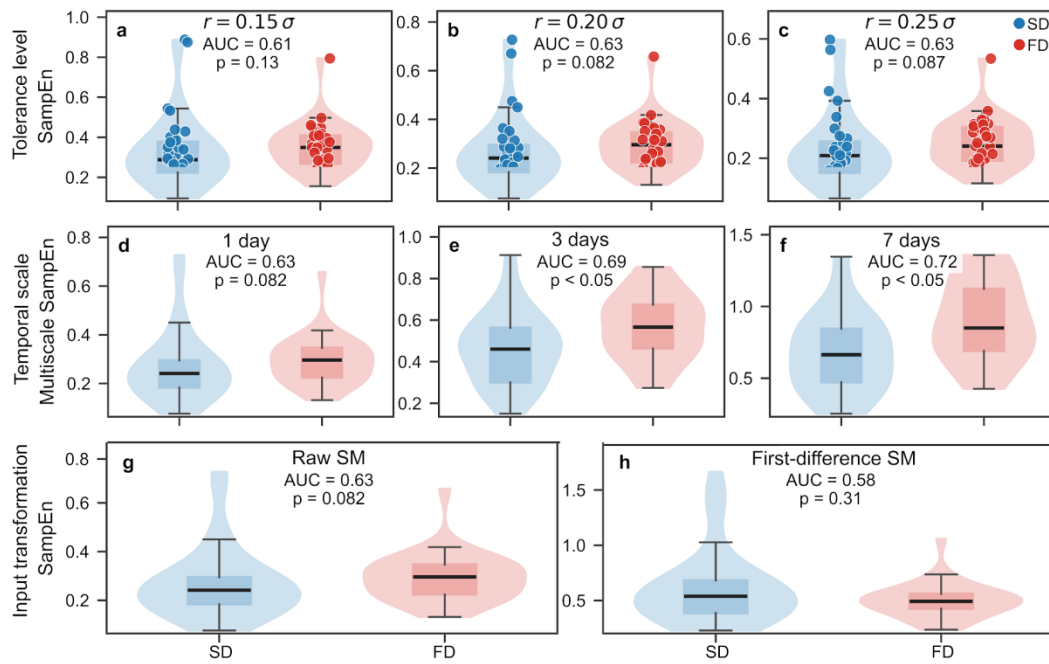


Fig. S16 Robustness of SampEn in discriminating FD and SD regimes under alternative settings. (a-c) Tolerance $r = 0.15, 0.20$, and $0.25 \times \text{SD}$. (d-f) Multiscale SampEn at 1-, 3-, and 7-day scales. (g-h) Raw vs. first-difference SM series. AUC and p-values are from Mann-Whitney U tests.

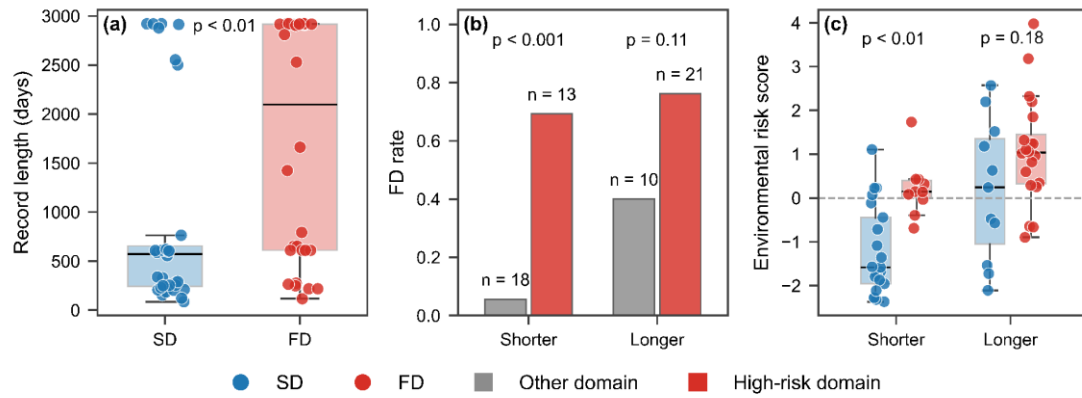


Fig. S17 Length-stratified validation of rainfall-soil risk effects, splitting sites at the median record length. (a) Record length by regime. (b) FD rate inside vs. outside the FD regime domain, by length stratum. (c) FDI by regime, by length stratum.

Supplementary References

Carsel, R.F., Parrish, R.S., 1988. Developing joint probability distributions of soil water retention characteristics. *Water Resources Research* 24(5), 755–769. <https://doi.org/10.1029/WR024i005p00755>

Mualem, Y., 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research* 12(3), 513–522. <https://doi.org/10.1029/WR012i003p00513>

van Genuchten, M.Th., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal* 44(5), 892–898. <https://doi.org/10.2136/sssaj1980.03615995004400050002x>