Supplement of Earth Syst. Sci. Data, 15, 395–409, 2023
https://doi.org/10.5194/essd-15-395-2023-supplement

Open Access Earth System Science Data

Supplement of

ChinaCropSM1 km: a fine 1 km daily soil moisture dataset for dryland wheat and maize across China during 1993–2018

Fei Cheng et al.

Correspondence to: Zhao Zhang (sunny_zhang@bnu.edu.cn)

Table S1 Means and medians of evaluation indices on spatial accuracy among ChinaCropSM1 km, RSSSM and ESA CCI SM, with better indices in bold.

Table S2 Means and medians of evaluation indices on temporal accuracy among ChinaCropSM1 km, RSSSM and ESA CCI SM, with better values in bold.

Table S3 Confusion matrix table in this study.

Table S4 Confusion matrix of irrigated validation based on the test dataset. Prediction categories are columns, while reference categories are rows.

Table S5 The accuracy comparison between with irrigation module (in bold) and without it.

Figure S1 The locations of all meteorological stations in China.

Figure S2 An overview of the workflow to develop an irrigation module to forecast soil moisture based on RF models.

Figure S3 Correlation coefficient of each factor and soil moisture in wheat-planted land, *, ** and *** for significance levels at $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.

Figure S4 Correlation coefficient of each factor and soil moisture in maize-planted land; *, ** and *** indicate the same as those in Figure S1

Figure S5 The accuracy (negative mean of absolute error) of the RF models with all selected hyperparameters.

Figure S6 Training and testing samples for temporal pattern comparison between ChinaCropSM1 km and in situ soil moisture observations.

Figure S7 Comparison between the predicted soil moisture (ChinaCropSM1 km) and in situ samples by crops and depths (cm) in the training set. (a) $wheat_{0–10}$, (b) $wheat_{10–20}$, (c) $maize_{0–10}$ and (d) $maize_{10–20}$. The red lines are the trend lines, the color bar indicates the point density, and the black lines represent the 1:1 lines.

Table S1 Means and medians of evaluation indices on spatial accuracy among ChinaCropSM1 km, RSSSM and ESA CCI SM, with better values in bold.

INDEX	*r*			RMSE			bias			ubRMSE		
	$maize_{0-10}$	RSSSM	ESA CCI SM	$maize_{0-10}$	RSSSM	ESA CCI SM	$maize_{0-10}$	RSSSM	ESA CCI SM	$maize_{0-10}$	RSSSM	ESA CCI SM
Mean	**0.947**	0.376	0.303	**0.027**	0.167	0.121	**0.0006**	–0.138	–0.067	**0.026**	0.085	0.092
Median	**0.946**	0.458	0.295	**0.030**	0.166	0.120	**0.0006**	–0.133	–0.075	**0.029**	0.084	0.092
	$maize_{10-20}$	RSSSM	ESA CCI SM	$maize_{10-20}$	RSSSM	ESA CCI SM	$maize_{10-20}$	RSSSM	ESA CCI SM	$maize_{10-20}$	RSSSM	ESA CCI SM
Mean	**0.957**			**0.032**			**0.001**			**0.020**		
Median	**0.958**			**0.035**			**0.0005**			**0.024**		
	$wheat_{0-10}$	RSSSM	ESA CCI SM	$wheat_{0-10}$	RSSSM	ESA CCI SM	$wheat_{0-10}$	RSSSM	ESA CCI SM	$wheat_{0-10}$	RSSSM	ESA CCI SM
Mean	**0.931**	0.306	0.184	**0.051**	0.181	0.111	**0.0006**	–0.153	–0.055	**0.031**	0.089	0.092
Median	**0.922**	0.324	0.193	**0.051**	0.183	0.112	**0.0014**	–0.155	–0.053	**0.034**	0.095	0.094
	$wheat_{10-20}$	RSSSM	ESA CCI SM	$wheat_{10-20}$	RSSSM	ESA CCI SM	$wheat_{10-20}$	RSSSM	ESA CCI SM	$wheat_{10-20}$	RSSSM	ESA CCI SM
Mean	**0.947**			**0.026**			**–0.0008**			**0.025**		
Median	**0.946**			**0.027**			**0.0003**			**0.026**		

Note: *r*: Pearson correlation coefficient; RMSE: root mean square error (m^3m^{-3}); bias (m^3m^{-3}); ubRMSE: unbiased RMSE (m^3m^{-3}); $wheat_{0–10}$: the 1 km-gridded daily soil moisture dataset for wheat-planted land at 0–10 cm depth; $wheat_{10–20}$: the 1 km-gridded daily soil moisture dataset for wheat-planted land at 10–20 cm depth, $maize_{0–10}$: the 1 km-gridded daily soil moisture dataset for maize-planted land at 0–10 cm depth; $maize_{10–20}$: the 1 km-gridded daily soil moisture dataset for maize-planted land at 10–20 cm depth; ChinaCroplandSM1 km: the 1 km soil moisture dataset for dry croplands in China; RSSSM: the global remote-sensing-based surface soil moisture dataset; ESA CCI SM: the European Space Agency Climate Change Initiative soil moisture product.

Table S2 Means and medians of evaluation indices on temporal accuracy among ChinaCropSM1 km, RSSSM and ESA CCI SM, with better performance in bold.

INDEX	*r*			RMSE			bias			ubRMSE		
	$maize_{0\text{-}10}$	RSSSM	ESA CCI SM	$maize_{0\text{-}10}$	RSSSM	ESA CCI SM	$maize_{0\text{-}10}$	RSSSM	ESA CCI SM	$maize_{0\text{-}10}$	RSSSM	ESA CCI SM
Mean	**0.830**	0.307	0.380	**0.036**	0.156	0.110	**–0.0009**	–0.137	–0.075	**0.033**	0.058	0.054
Median	**0.886**	0.399	0.484	**0.033**	0.148	0.096	**0.0005**	–0.134	–0.074	**0.031**	0.058	0.052
	$maize_{10\text{-}20}$	RSSSM	ESA CCI SM	$maize_{10\text{-}20}$	RSSSM	ESA CCI SM	$maize_{10\text{-}20}$	RSSSM	ESA CCI SM	$maize_{10\text{-}20}$	RSSSM	ESA CCI SM
Mean	**0.833**			**0.027**			**–0.0004**			**0.026**		
Median	**0.895**			**0.027**			**0.0008**			**0.025**		
	$wheat_{0\text{-}10}$	RSSSM	ESA CCI SM	$wheat_{0\text{-}10}$	RSSSM	ESA CCI SM	$wheat_{0\text{-}10}$	RSSSM	ESA CCI SM	$wheat_{0\text{-}10}$	RSSSM	ESA CCI SM
Mean	**0.821**	0.252	0.397	**0.038**	0.163	0.102	**0.002**	–0.143	–0.059	**0.034**	0.059	0.054
Median	**0.853**	0.310	0.486	**0.037**	0.157	0.085	**0.002**	–0.141	–0.057	**0.033**	0.058	0.054
	$wheat_{10\text{-}20}$	RSSSM	ESA CCI SM	$wheat_{10\text{-}20}$	RSSSM	ESA CCI SM	$wheat_{10\text{-}20}$	RSSSM	ESA CCI SM	$wheat_{10\text{-}20}$	RSSSM	ESA CCI SM
Mean	**0.841**			**0.030**			**0.0009**			**0.028**		
Median	**0.875**			**0.028**			**0.0007**			**0.027**		

Note: *r*: Pearson correlation coefficient; RMSE: root mean square error (m^3m^{-3}); bias (m^3m^{-3}); ubRMSE: unbiased RMSE (m^3m^{-3}); $wheat_{0–10}$: the 1 km-gridded daily soil moisture dataset for wheat-planted land at 0–10 cm depth; $wheat_{10–20}$: the 1 km-gridded daily soil moisture dataset for wheat-planted land at 10–20 cm depth, $maize_{0–10}$: the 1 km-gridded daily soil moisture dataset for maize-planted land at 0–10 cm depth; $maize_{10–20}$: the 1 km-gridded daily soil moisture dataset for maize-planted land at 10–20 cm depth; ChinaCroplandSM1 km: the 1 km soil moisture dataset for dry croplands in China; RSSSM: the global remote-sensing-based surface soil moisture dataset; ESA CCI SM: the European Space Agency Climate Change Initiative soil moisture product.

Table S3 Confusion matrix table in this study.

		Class	
		Irrigated	Non
Reference	Irrigated	TP	FN
	Non	FP	TN

Table S4 Confusion matrix of irrigated validation based on the test dataset. Prediction categories are columns, while reference categories are rows.

ChinaCropSM1 km	Class	Irrigated	Non	Total	Accuracy	PA	UA	AUC
	Irrigated	1633	395	2028	0.85	0.82	0.81	0.84
$wheat_{0–10}$	Non	365	2744	3109				
	Total	1998	3139					
	Irrigated	1583	446	2029	0.84	0.81	0.78	0.83
$wheat_{10–20}$	Non	365	2749	3114				
	Total	1948	3195					
	Irrigated	915	310	1225	0.86	0.85	0.75	0.84
$maize_{0–10}$	Non	167	2030	2197				
	Total	1082	2340					
	Irrigated	875	321	1196	0.86	0.83	0.73	0.83
$maize_{10–20}$	Non	175	2052	2227				
	Total	1050	2373					

Table S5 The accuracy comparison between with irrigation module (in bold) and without it.

ChinaCropSM1 km	BIAS		r		RMSE		ubRMSE	
$wheat_{0–10}$	**−0.0011**	−0.0019	**0.860**	0.801	**0.037**	0.044	**0.037**	0.044
$wheat_{10–20}$	**−0.0002**	−0.0006	**0.895**	0.838	**0.031**	0.039	**0.031**	0.039
$maize_{0–10}$	**0.0009**	0.0007	**0.861**	0.798	**0.036**	0.043	**0.036**	0.043
$maize_{10–20}$	**0.0003**	−0.0001	**0.894**	0.812	**0.029**	0.038	**0.029**	0.038

Figure S1. The locations of all meteorological stations in China.

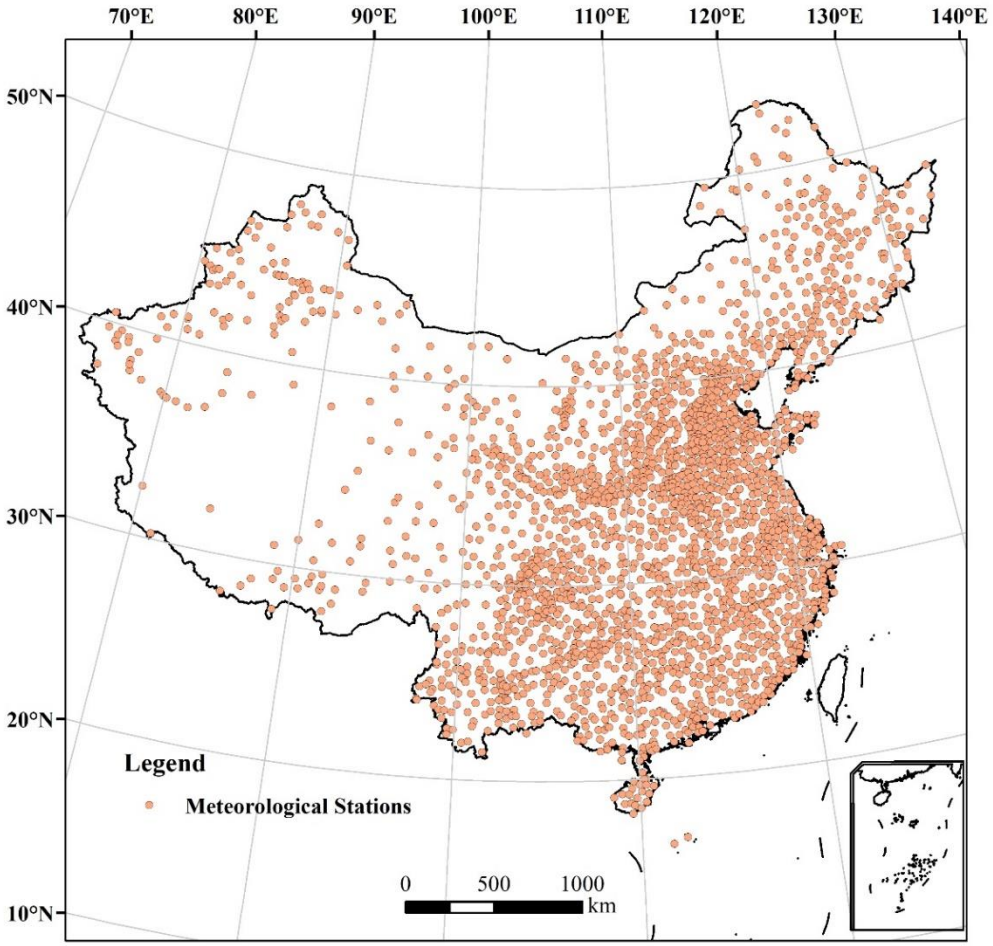
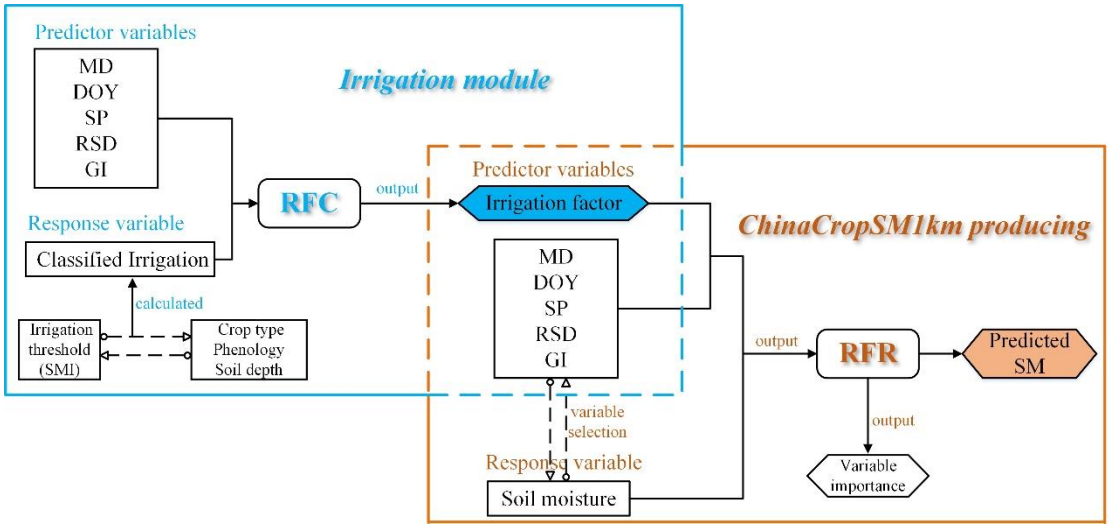
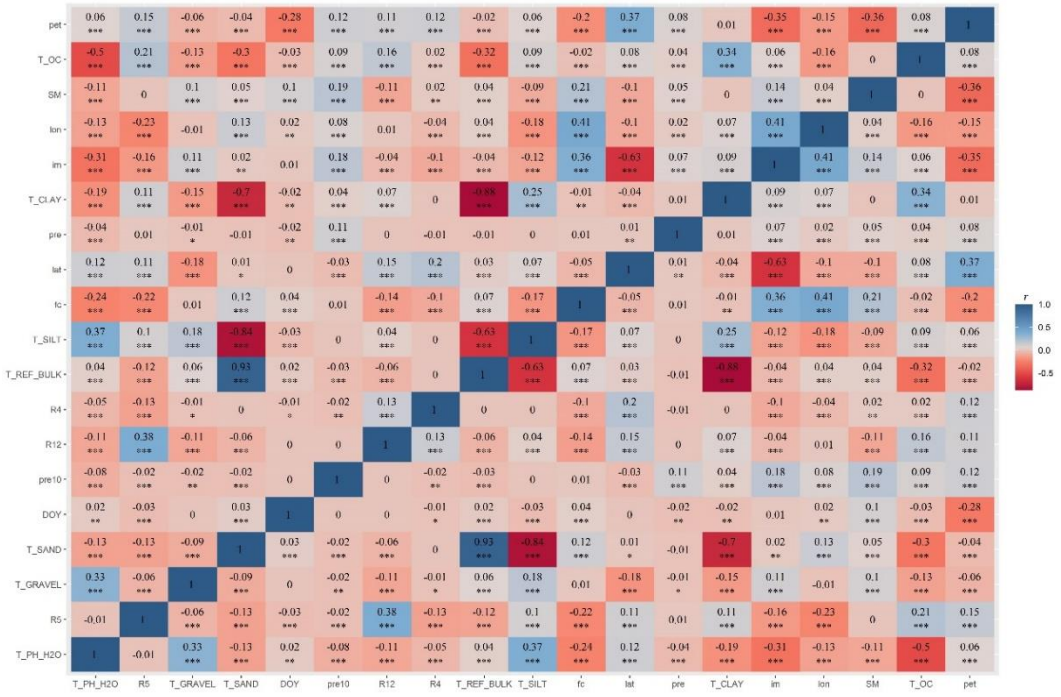


Figure S2. An overview of the workflow to develop an irrigation module to forecast soil moisture based on RF models.



Note: SM: soil moisture; RFC: random forest classification; RFR: random forest regression; SMI: evaluation index of relative soil moisture to determine when irrigation is applied; MD: meteorological data; DOY: day of year; SP: soil properties; RSD: remote sensing data; GI: geographical information.

Figure S3. Correlation coefficient of each factor and soil moisture in wheat-planted land, *, ** and * for significance levels at $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.**



Note: *r*: Pearson correlation coefficient; SM: soil moisture; CIR: classified irrigation; pre10: anteaccumulated precipitation over ten days; fc: field capacity; DOY: day of year; lon: longitude; pre: daily precipitation; im: moisture index; lat: latitude; pet: reference evapotranspiration; R4: river network vector I; R5: river network vector II; R12: river network vector III; REF_BULK: soil bulk density; PH_H2O: hydrogen ion concentration; GRAVEL: volume percentage of crushed stone; T: the topsoil layer.

Figure S4. Correlation coefficient of each factor and soil moisture in maize-planted land; *, ** and * indicate the same as those in Figure S1.**

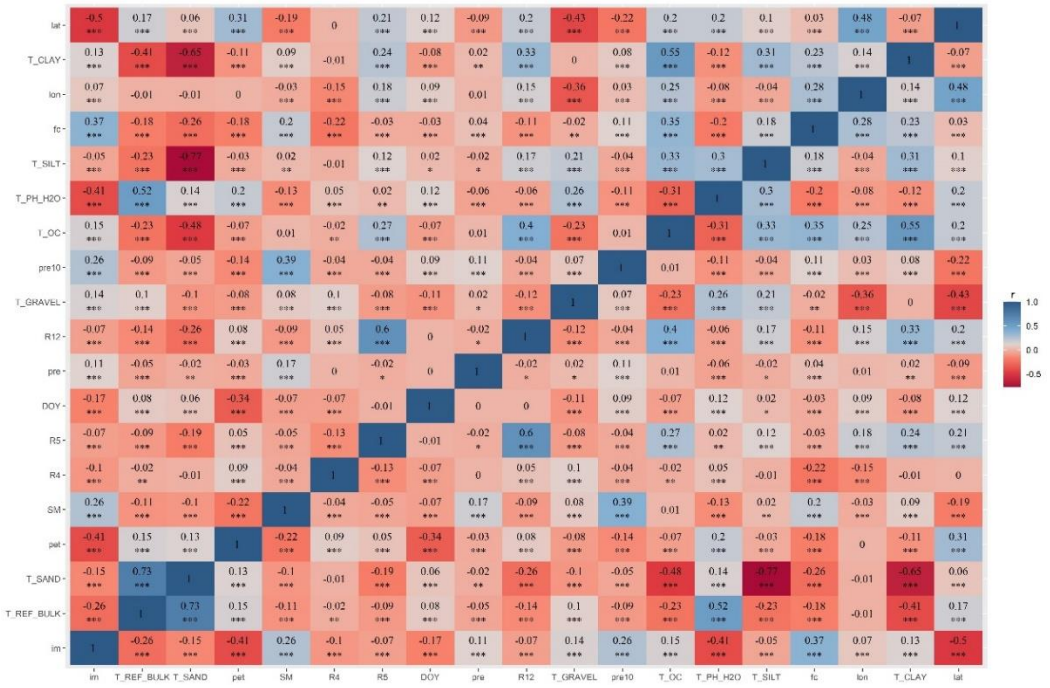


Figure S5. The accuracy (negative mean of absolute error) of the RF models with all selected hyperparameters.

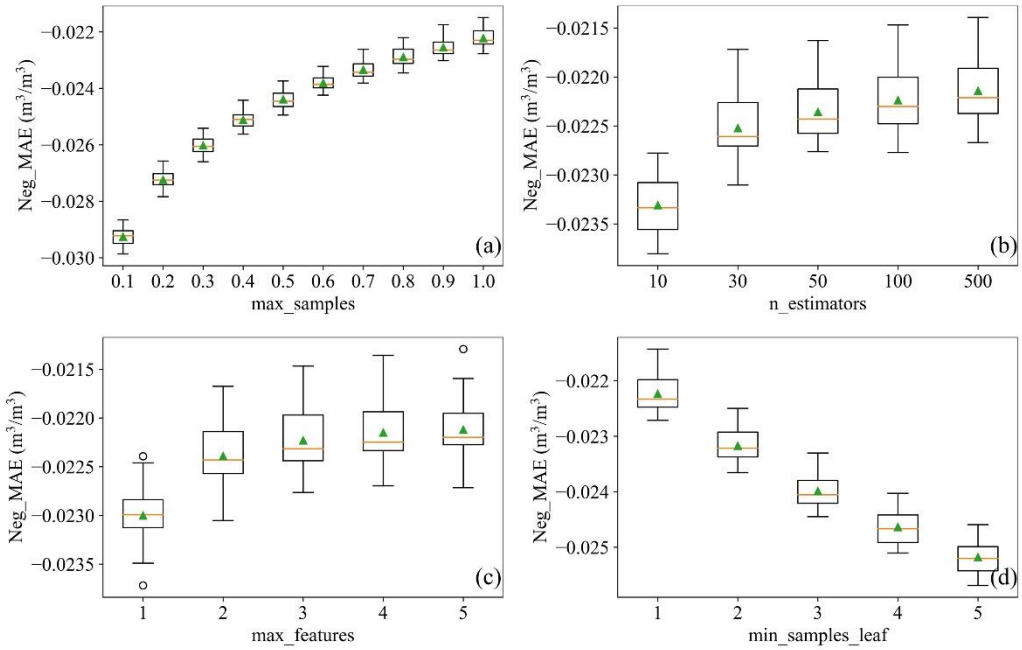
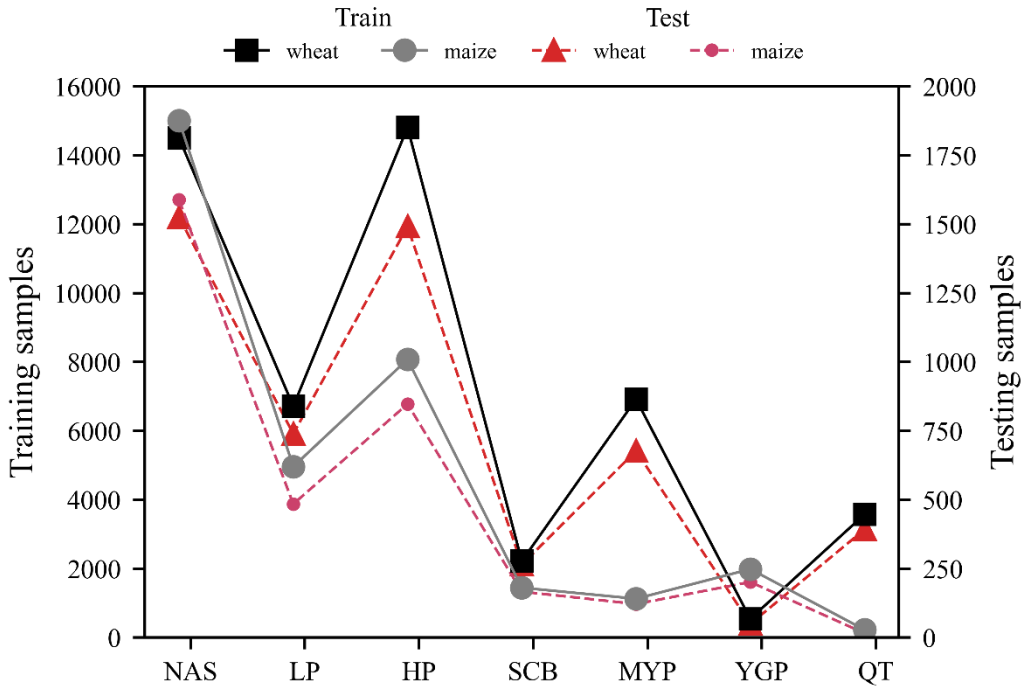


Figure S6. Training and testing samples for temporal pattern comparison between ChinaCropSM1 km and in situ soil moisture observations.



Note: NAS: Northern Arid and Semiarid region; LP: Loess Plateau; HP: Huang–Huai–Hai Plain; SCB: SiChuan Basin; MYP: Middle–lower Yangtze Plain; YGP: Yunnan–Guizhou Plateau and southern China; QT: Qinghai–Tibet region.

Figure S7. Comparison between the predicted soil moisture (ChinaCropSM1 km) and in situ samples by crops and depths (cm) in the training set. (a) wheat$_{0–10}$, (b) wheat$_{10–20}$, (c) maize$_{0–10}$ and (d) maize$_{10–20}$. The red lines are the trend lines, the color bar indicates the point density, and the black lines represent the 1:1 lines.