



## 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)

The copyright of individual parts of the supplement might differ from the article licence.

**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.

20 **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<sub>0-10</sub>, (b) wheat<sub>10-20</sub>, (c) maize<sub>0-10</sub> and (d) maize<sub>10-20</sub>. 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, RSSSMand ESA CCI SM, with better values in bold.

INDEX		r			RMSE				bias				ubRMSE	3
	maize <sub>0-10</sub>	RSSSM	ESA CCI SM	maize <sub>0-10</sub>	RSSSM	ESA CCI SM	-	maize <sub>0-10</sub>	RSSSM	ESA CCI SM	maiz	e <sub>0-10</sub>	RSSSM	ESA CCI SM
Mean	0.947	0.376	0.303	0.027	0.167	0.121	-	0.0006	-0.138	-0.067	0.0	26	0.085	0.092
Median	0.946	0.458	0.295	0.030	0.166	0.120		0.0006	-0.133	-0.075	0.0	29	0.084	0.092
	maize <sub>10-20</sub>	RSSSM	ESA CCI SM	maize <sub>10-20</sub>	RSSSM	ESA CCI SM	-	maize <sub>10-20</sub>	RSSSM	ESA CCI SM	maiz	e <sub>10-20</sub>	RSSSM	ESA CCI SM
Mean	0.957			0.032			-	0.001			0.0	20		
Median	0.958			0.035	-		_	0.0005	-		0.0	24	-	
	wheat <sub>0-10</sub>	RSSSM	ESA CCI SM	wheat <sub>0-10</sub>	RSSSM	ESA CCI SM		wheat <sub>0-10</sub>	RSSSM	ESA CCI SM	whea	at <sub>0-10</sub>	RSSSM	ESA CCI SM
Mean	0.931	0.306	0.184	0.051	0.181	0.111	-	0.0006	-0.153	-0.055	0.0	31	0.089	0.092
Median	0.922	0.324	0.193	0.051	0.183	0.112		0.0014	-0.155	-0.053	0.0	34	0.095	0.094
	wheat10-20	RSSSM	ESA CCI SM	wheat <sub>10-20</sub>	RSSSM	ESA CCI SM	-	wheat <sub>10-20</sub>	RSSSM	ESA CCI SM	whea	ıt <sub>10-20</sub>	RSSSM	ESA CCI SM
Mean	0.947			0.026			-	-0.0008			0.0	25		
Median	0.946			0.027	-			0.0003			0.0	26	-	

Note: *r*: Pearson correlation coefficient; RMSE: root mean square error (m<sup>3</sup>m<sup>-3</sup>); bias (m<sup>3</sup>m<sup>-3</sup>); ubRMSE: unbiased RMSE (m<sup>3</sup>m<sup>-3</sup>); wheat<sub>0-10</sub>: the 1 km-gridded daily soil moisture dataset for wheat-planted land at 0–10 cm depth; wheat<sub>10-20</sub>: the 1 km-gridded daily soil moisture dataset for wheat-planted land at 10–20 cm depth, maize<sub>0-10</sub>: the 1 km-gridded daily soil moisture dataset for maize-planted land at 0–10 cm depth; maize<sub>10-20</sub>: the 1 km-gridded daily soil moisture dataset for maize-planted land at 0–10 cm depth; maize<sub>10-20</sub>: the 1 km-gridded daily soil moisture dataset for maize-planted land at 0–10 cm depth; maize<sub>10-20</sub>: 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				ubRMS	E
	maize <sub>0-10</sub>	RSSSM	ESA CCI SM	-	maize <sub>0-10</sub>	RSSSM	ESA CCI SM	•	maize <sub>0-10</sub>	RSSSM	ESA CCI SM		maize <sub>0-10</sub>	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 <sub>10-20</sub>	RSSSM	ESA CCI SM	-	maize <sub>10-20</sub>	RSSSM	ESA CCI SM		maize <sub>10-20</sub>	RSSSM	ESA CCI SM		maize <sub>10-20</sub>	RSSSM	ESA CCI SM
Mean	0.833			-	0.027			•	-0.0004				0.026		
Median	0.895				0.027				0.0008				0.025		
	wheat <sub>0-10</sub>	RSSSM	ESA CCI SM	-	wheat <sub>0-10</sub>	RSSSM	ESA CCI SM		wheat <sub>0-10</sub>	RSSSM	ESA CCI SM		wheat <sub>0-10</sub>	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
	wheat10-20	RSSSM	ESA CCI SM	-	wheat10-20	RSSSM	ESA CCI SM		wheat10-20	RSSSM	ESA CCI SM		wheat10-20	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<sup>3</sup>m<sup>-3</sup>); bias (m<sup>3</sup>m<sup>-3</sup>); ubRMSE: unbiased RMSE (m<sup>3</sup>m<sup>-3</sup>); wheat<sub>0-10</sub>: the 1 km-gridded daily soil moisture dataset for wheat-planted land at 0–10 cm depth; wheat<sub>10-20</sub>: the 1 km-gridded daily soil moisture dataset for wheat-planted land at 10–20 cm depth, maize<sub>0-10</sub>: the 1 km-gridded daily soil moisture dataset for maize-planted land at 0–10 cm depth; maize<sub>10-20</sub>: the 1 km-gridded daily soil moisture dataset for maize-planted land at 0–10 cm depth; maize<sub>10-20</sub>: the 1 km-gridded daily soil moisture dataset for maize-planted land at 0–10 cm depth; maize<sub>10-20</sub>: 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; ESA CCI SM: the European Space Agency Climate Change Initiative soil moisture product.

## Table S3 Confusion matrix table in this study.

		Class				
		Irrigated	Non			
Defense	Irrigated	TP	FN			
Reference	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 <sub>0-10</sub>	Non	365	2744	3109				
	Total	1998	3139					
	Irrigated	1583	446	2029	0.84	0.81	0.78	0.83
wheat <sub>10-20</sub>	Non	365	2749	3114				
	Total	1948	3195					
	Irrigated	915	310	1225	0.86	0.85	0.75	0.84
maize <sub>0-10</sub>	Non	167	2030	2197				
	Total	1082	2340					
	Irrigated	875	321	1196	0.86	0.83	0.73	0.83
maize <sub>10-20</sub>	Non	175	2052	2227				
	Total	1050	2373					

ChinaCropSM1 km	BIAS		r		RN	ISE	ubRMSE		
wheat <sub>0-10</sub>	-0.0011	-0.0019	0.860	0.801	0.037	0.044	0.037	0.044	
wheat <sub>10-20</sub>	-0.0002	-0.0006	0.895	0.838	0.031	0.039	0.031	0.039	
maize <sub>0-10</sub>	0.0009	0.0007	0.861	0.798	0.036	0.043	0.036	0.043	
maize <sub>10-20</sub>	0.0003	-0.0001	0.894	0.812	0.029	0.038	0.029	0.038	

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





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.

0.15 0.11 0.12 -0.15 0.06 0.12 0.08 0.08 -0.06 0.01 -0.2 0.34 \*\*\* -0.16 0.21 -0.13 -0.3 -0.03 \*\*\* 0.09 0.16 0.02 -0.32 0.09 -0.02 80.0 0.04 \*\*\* 0.06 0.08 T\_OC 0 0.05 0.19 -0.11 0.02 -0.09 0.21 -0.1 0.05 \*\*\* 0.14 0.04 -0.11 0.1 \*\*\* 0.1 0.04 SM 0 0 0 \*\*\* -0.13 -0.23 0.13 0.02 0.08 -0.04 0.04 -0.18 0.41 -0.1 0.02 0.07 0.41 0.04 -0.16 -0.15 lon -0.01 0.01 -0.04 -0.12 0.07 -0.31 -0.16 0.11 0.02 0.18 -0.1 \*\*\* -0.04 0.36 0.09 0.41 0.14 0.06 -0.35 0.01 0.11 0.15 0.04 0.07 0.25 -0.04 \*\*\* 0.34 \*\*\* -0.19 \*\*\* -0.02 -0.01 \*\* 0.07 0.01 0.09 \*\*\* 0 0 0.01 T CLAY 0.11 -0.04 -0.01 -0.02 0.01 0.07 0.02 0.05 0.04 0.08 \*\*\* 0.01 -0.01 0 -0.01 -0.01 0.01 0.01 0 -0.1 0.37 \*\*\* 0.12 0.11 -0.18 0.01 -0.03 \*\*\* 0.15 0,2 \*\*\* 0.03 0.07 -0.05 0.01 -0.04 -0.1 0.08 \*\*\* lat 0 0,41 \*\*\* 0,21 -0.02 -0,14 -0,1 \*\*\* 0,07 -0.17 0,24 0,04 -0.01 0,12 -0,05 0.01 0.01 0.01 0.5 0.0 -0.5 0.06 \*\*\* -0.02 \*\*\* 0.37 0.1 0.18 -0.03 0.04 0.07 0.25 0,18 -0.09 0.09 T\_SILT -0 0 -0,17 0 -0.12 -0.03 \*\*\* 0,03 0,04 \*\*\* 0.06 0.02 -0.06 0.07 0,04 ==== 0,04 \*\*\* 0,04 \*\*\* -0.32 0 T REF BULK. -0.01 -0.1 0.12 -0.05 \*\*\* -0.13 -0.01 -0.01 -0.02 0.13 0.2 -0,1 -0.04 0.02 0.02 R4 -U 0 -0.01 0 0.13 -0.14 0.15 -0.04 -0.11 0.11 -0.11 0.38 -0.11 -0.06 -0.06 0.07 0.16 R12 -0 0.04 0.01 0 0 -0.08 -0.02 -0.02 -0.02 -0.02 -0.03 -0.03 0.11 0.04 0.18 \*\*\* 0.08 0.19 0.09 0.12 pre10 0 0 0,01 0.02 0.03 0.02 0.04 -0.02 0.1 -0.03 -0.28 0.02 -0.03 -0.01 -0.03 -0.02 0.01 DOY 0 0 0 0 -0.13 -0.13 -0.02 -0.06 0.12 0.01 0.02 0.13 0.05 -0.04 -0.3 0.03 T\_SAND -0,01 0.33 -0.06 -0.02 -0.11 -0.01 -0.18 -0.01 0.11 0.1 -0.13 -0.06 0.09 T GRAVEL 0 0.01 -0.01 -0.22 \*\*\* 0.11 0.23 \*\*\* 0.15 •0.13 -0.03 \*\*\* -0.02 \*\*\* 0.38 -0.13 -0.12 \*\*\* 0.1 0.11 -0.16 \*\*\* 0.21 0.06 0 -0,01 0,01 R5 · 0.12 0.33 -0.13 0.02 -0.08 -0.05 0.04 0.37 0.24 -0.04 0.31 -0.13 -0.11 0.06 0.19 T\_PH\_H -0.01 R4

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.

0.17 0.06 0.31 0.21 0.12 0.2 0.2 0.2 0.1 0.03 0.48 -0.19 -0.09 -0.22 -0.07 lat -0 -0.12 0.23 0.24 0.33 0.31 0.14 0.09 -0.08 0.02 0.08 0.55 0.13 -0.11 -0.01 -0.0 T\_CLAY -0 -0.08 \*\*\* -0.2 \*\*\* 0.14 \*\*\* 0.23 \*\*\* 0.28 0.48 0.07 -0.03 -0,15 0.18 0,09 0.15 -0,36 0.03 0.25 -0.04 -0.01 -0.01 0 0.01 -0.03 -0.26 -0.18 -0.22 -0.03 -0.11 -0.02 0.11 \*\*\* 0.35 0.18 \*\*\* 0.37 -0.18 0.2 0.04 0.28 0.03 fc--0.03 0.3 0.31 0.02 0.12 \*\*\* -0.04 \*\*\* 0,02 \* -0,02 # 0.17 -0,04 \*\*\* 0,33 0.1 \*\*\* -0.05 -0.23 0.21 T\_SILT -+0.01 0.18 3## 0.2 \*\*\* -0.07 0.52 0.14 \*\*\* -0.13 0.05 0.02 0.12 -0.06 -0.06 0.26 -0.11 0.3 -0.08 -0.12 0.2 T\_PH\_H2O 0.2 •••• •0.22 •••• 0.27 0.4 0.33 \*\*\* 0.25 0.35 0.55 -0.02 -0.07 0.15 -0.23 0.01 0.01 T\_0C-0.01 -0.04 \*\*\* -0.14 -0.04 \*\*\* -0.04 \*\*\* 0.03 0.08 0.07 -0.11 0.11 0.26 -0.09 -0.05 0.39 0.09 0.11 0,01 pre10-0,1 0.02 0,14 0,1 -0.1 \*\*\* -0.08 0.08 -0.08 0,11 -0,12 0,07 8.9# -0.23 0.26 0.21 -0.02 0.43 0 T GRAVEL 1.0 0.5 0.0 -0.5 -0.06 0.6 \*\*\* 0.4 0.17 \*\*\* 0.33 \*\*\* -0.26 0.2 \*\*\* -0.07 -0.14 \*\*\* 0.08 \*\*\* -0.09 0.05 -0.02 -0.04 \*\*\* -0.11 \*\*\* -0.12 0.15 R12-0 -0.09 0.11 -0.05 -0.02 -0.03 \*\*\* 0.17 0.02 0,02 \* 0.11 -0.06 -0.02 0.04 0.02 -0,02 pre 0 0 0.01 0.01 -0.07 -0.07 0.12 0.12 -0.17 0.08 0.06 -0.07 -0.11 0.09 \*\*\* 0.02 -0.03 0.09 -0.08 DOY--0.01 0 0 -0.08 0.02 \*\* 0.05 -0.05 -0.13 -0.02 -0.04 0.27 0.12 0.18 0.24 0.21 -0.07 -0.09 -0.19 -0.03 R5 --0.01 -0.1 -0.02 0.09 \*\*\* -0.04 -0.07 0.1 -0.04 -0.02 0.05 -0.22 -0.15 -0.13 -0.01 -0.01 R4 -0 -0.01 0 -0.13 \*\*\* -0.07 0.08 0.39 \*\*\* -0.14 \*\*\* -0.1 0.17 0.2 -0.19 0.26 -0.11 -0.05 -0.09 0.02 -0.03 0.09 -0.04 \*\*\* SM-0.01 -0.11 94 0.09 0.05 0.08 -0.08 -0.07 0.2 0.31 0.15 0.13 0.34 -0.03 -0.03 -0.18 0 -0.15 0.13 -0.1 -0.19 0.06 -0.02 -0.26 -0.1 -0.05 0.14 \*\*\* 0.06 -0.26 -0,01 T SAND -0,01 -0.02 -0.09 0++ -0.14 0.1 0.52 0.17 0.15 -0.11 -0.18 -0.09 0.08 -0.05 -0.26 T\_REF\_BULK --0.01 0.26 -0.1 -0.07 -0.17 0.11 0.07 0.14 0.26 0.15 -0.05 0.37 0.07 0.13 -0.15 IM T REF BULK T SAND pet SN R т юс T PH H20 T\_SILT T CLAY

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<sub>0-10</sub>, (b) wheat<sub>10-20</sub>, (c) maize<sub>0-10</sub> and (d) maize<sub>10-20</sub>. The red lines are the trend lines, the color bar indicates the point density, and the black lines represent the 1:1 lines.

