A 1-km daily surface soil moisture dataset of enhanced coverage under all-weather conditions over China in 2003-2019

Peilin Song[1,4], Yongqiang Zhang[1]*, Jianping Guo[2]*, Jiancheng Shi[3], Tianjie Zhao[4], Bing Tong[2]

[1] Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, The Chinese Academy of Sciences, Beijing 100101, China

[2] State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China

[3] National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

[4] State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences. Beijing 100101, China

Correspondence to: Yongqiang Zhang (zhangyq@igsnrr.ac.cn); Jianping Guo (jpguo@cma.gov.cn)

Abstract:

Surface soil moisture (SSM) is crucial for understanding the hydrological process of our earth surface. Passive microwave (PM) technique has long been the primary tool for estimating global SSM from the view of satellite, while the coarse resolution (usually >~10 km) of PM observations hampers its applications at finer scales. Although quantitative studies have been proposed for downscaling satellite PM-based SSM, very few products have been available to public that meet the qualification of 1-km resolution and daily revisit cycles under all-weather conditions. In this study, we developed one such SSM product in China with all these characteristics. The product was generated through downscaling the AMSR-E/AMSR-2 based SSM at 36-km, covering all on-orbit time of the two radiometers during 2003-2019. MODIS optical reflectance data and daily thermal infrared land surface temperature (LST) that had been gap-filled for cloudy conditions were the primary data inputs of the downscaling model, so that the "all-weather" quality was achieved for the 1-km SSM. Daily images from this developed SSM product have quasi-complete coverage over the country during April-September. For other months, the national coverage percentage of the developed product is also greatly improved against the original daily PM observations, through a specifically developed sub-model for filling the gap between seams of neighboring PM swaths during the downscaling procedure. The product is well compared against *in situ* soil moisture measurements from 2000+ meteorological stations, indicated by station averages of the unbiased RMSD ranging from 0.052 vol/vol to 0.059 vol/vol. Moreover, the evaluation results also show that the developed

43   product outperforms the SMAP-Sentinel (Active-Passive microwave) combined SSM

44   product at 1-km, with a correlation coefficient of 0.55 achieved against that of 0.40 for

45   the latter product. This indicates the new product has great potential to be used for

46   hydrological community, agricultural industry, water resource and environment

47   management.

1. Introduction

49   Surface soil moisture (SSM) is one of the most important variables that dominate

50   the mass and energy cycles of earth surface system (Entekhabi et al., 2010b). Satellite-

51   based SSM datasets of sufficiently fine spatio-temporal resolutions over large-scale

52   areas have significant implication on improved investigations at various research fields

53   including hydrological signature identification (Zhou et al., 2021; Jung et al., 2010),

54   agricultural yield production estimation (Ines et al., 2013; Pan et al., 2019),

55   drought/waterlogging monitoring and warning (Vergopolan et al., 2021; Den Besten et

56   al., 2021; Jing and Zhang, 2010), as well as weather prediction and future climate

57   analysis (Koster et al., 2010; Jeffrey et al., 2001). Microwave bands with centimeter-

58   level or longer wavelengths (X-band, C-band, and L-band) are currently identified as

59   the primary band channels suitable for SSM observations from view of satellite, due to

60   their high penetration capabilities through cloud layers and vegetation canopies. In

61   terms of sensor types, microwave SSM detection includes passive microwave

62   (radiometer-based) techniques and active microwave (radar, scatterometer) techniques.

63   Satellite-based passive microwave (PM) radiometers, e.g. the Soil Moisture Active

64 Passive (SMAP), the Soil Moisture and Ocean Salinity (SMOS), and the Advance

65 Microwave Scanning Radiometer-2 (AMSR-2), can obtain SSM observations at a

66 revisit interval of 1-3 days, with relatively poor native spatial resolutions of tens of

67 kilometers. Active microwave (AM) such as radar can achieve kilometer-level and even

68 finer resolution of observations targeting at the earth surface. However, this usually

69 sacrifices the swath width of radar configuration, because of which, most satellite-based

70 synthetic aperture radars (SAR) have an obviously longer global revisit cycle (usually

71 longer than 5 days, e.g. Sentinel-1 SAR data) than the typical radiometers. Moreover,

72 AM radar backscatter signals are extremely sensitive to speckle noise (Entekhabi et al.,

73 2016), as well as influence from soil roughness, vegetation canopy structure and water

74 content (Piles et al., 2009). All above influential factors have seriously impeded the use

75 of AM radar techniques or combination of passive/active microwave datasets for

76 producing high spatial resolution SSM products with a frequent revisit.

77 Apart from microwave signals, solar reflectance or ground emission signals

78 originated from optical and infrared band domains also have the potential to reflect

79 SSM variation. Based on optical/infrared bands, however, SSM is typically estimated

80 based on indirect relationships through intermediate variables like soil evaporation

81 (Komatsu, 2003), vegetation condition (Zeng et al., 2004), or soil thermal inertia

82 (Verstraeten et al., 2006). To overcome the spatio-temporally instable performance on

83 SSM modelling that might be brought by such indirect relationships, they are typically

84 fused with the PM SSM datasets. In this manner, it can well reconcile the advantage of

85 PM observations with respect to its high sensitivity to SSM variation, as well as that of

86 optical/infrared observations with respect to its finer spatial resolutions at kilometer- or

87 even hectometer-levels. Such data fusion techniques are also known as downscaling

88 techniques of PM remote sensing SSM. Archetypal downscaling models include the

89 "universal triangle feature space (UTFS)"-based models (Chauhan et al., 2003; Choi

90 and Hur, 2012; Sanchez-Ruiz et al., 2014), the "DISaggregation based on a Physical

91 And Theoretical scale CHange (DISPACTH)" model (Merlin et al., 2010; Merlin et al.,

92 2005; Merlin et al., 2013; Merlin et al., 2008), and the "University of California, Los

93 Angeles (UCLA)" model (Peng et al., 2016). The physics of these models are mainly

94 based on the response of SSM variation to changes in soil evaporation or land surface

95 evapotranspiration. Another significant branch of such downscaling models are based

96 on the sensitivity of SSM to soil thermal inertia, which is quantified by diurnal LST

97 difference estimated from thermal-infrared wave bands (Fang and Lakshmi, 2013; Fang

98 et al., 2018).

99   Sabaghy et al. (2020) have shown that using optical and infrared data can achieve

100 finer-resolution SSM estimates which are better consistent with ground soil moisture

101 records, compared with using the radar datasets. Moreover, considering the short revisit

102 cycle (daily) of optical/infrared sensors onboard typical polar-orbit satellites, using

103 optical/infrared datasets to downscale PM SSM should be among the optimal methods

104 for obtaining SSM data with high spatio-temporal resolutions over national, continental,

105 or global scales. On the other hand, satellite remote sensing SSM products that are

106 characterized by 1-km resolution of daily revisit intervals and stable long time series

107 dating back to at least 15-20 years ago, are urgently required for accelerating the

108 development of various research fields, especially agriculture industry, water resources

109 management, and hydrological disaster monitoring (Sabaghy et al., 2020; Mendoza et

110 al., 2016). However, very seldom sets of such data products are publicly available to

111 the remote sensing research community because of the following drawbacks. First,

112 there is a serious lack of cloud-free optical/infrared imagery, which means the method

113 cannot deliver any SSM downscaling under cloudy/rainy weather. Second, most of the

114 above-mentioned optical/infrared-data-based downscaling methods were mainly

115 evaluated at regional or even smaller scales. This might raise concern on the

116 universality of those methods. For example, the DISPATCH method has been

117 recognized to be less effective in humid (energy-limited) regions compared with in arid

118 and semi-arid (water-limited) regions (Molero et al., 2016; Song et al., 2021; Zheng et

119 al., 2021). As far as the UTFS-based method is concerned, a poorer performance was

120 obtained compared to the DISPATCH in a typical water-limited region in North

121 America, according to the experiment conducted by Kim and Hogue (2012).

122 To improve the above-mentioned issues, we produced an all-weather daily SSM

123 data product at 1-km resolution all over China during 2003-2019, based on fusion of

124 multiple remote sensing techniques, including reconstruction of optical/infrared

125 observations under cloud as well as an improved PM SSM downscaling methodology

126 proposed in our previous study (Song et al., 2021). The potential significance of this

127 study includes

128 (i) to better serve and investigate the land surface hydrology processes and their

129 sophisticated interactions to human society at multi-scale (from national to regional)

130 resolutions in China because the country covers about 1/15 of the global terrestrial area

131 with about 1/5 of the world population, and

132 (ii) to provide a methodology framework that can inspire future studies on

133 generating similar SSM datasets all over the globe, based on the plentifulness of

134 resources on climate type, land covers, and topography in China.

135

2. Methods and Materials

2.1 Datasets

138 2.1.1 PM SSM data

139 Spatial downscaling of PM SSM is the fundamental theory for constructing the

140 target finer-resolution data product in this study. Therefore, the native retrieval

141 accuracy of the coarse-resolution PM SSM data product, based on which the

142 downscaling procedures are performed, is considerably crucial to the performance of

143 the downscaled data product (Busch et al., 2012; Im et al., 2016; Kim and Hogue, 2012).

144 Although the L-band PM brightness temperature (TB) observed by satellite missions

145 such as SMAP or SMOS are considered more suitable for SSM retrieval compared with

146 C- or X-band TB, both above missions started their space operations in the 2010s. This

147 means that to obtain downscaled SSM of longer historical periods, we still require to

148 rely on other C-/X-band-based radiometers which started their operations earlier than

149 SMAP and SMOS. An optimal satellite PM TB observation system dating back to

150 earlier years of this century is composed of the "Advanced Microwave Scanning

Radiometer of the Earth Observing System (AMSR-E)", together with its successor of

AMSR-2. AMSR-E operated during 2002-2011 onboard the Aqua satellite which is

governed by National Aeronautics and Space Administration (NASA), whilst AMSR-

2 is operating onboard the Global Change Observation Mission1-Water (GCOM-W1)

satellite developed by the Japan Aerospace Exploration Agency (JAXA) since 2012.

Several classical PM SSM retrieval algorithms have been applied to the afore-

mentioned "AMSR series (including AMSR-E and AMSR-2)" TB for generating long-

term global SSM products at 25 km (Table 1Table 1), including the JAXA algorithm

(Fujii et al., 2009; Koike et al., 2004), the "Land Parameter Retrieval Model (LPRM)"

algorithm (Song et al., 2019b; Meesters et al., 2005; Owe et al., 2001), and the

algorithm developed by the University of Montana (UMT) (Jones et al., 2009; Du et al.,

2016). A recent AMSR-based night-time SSM product during 2002-2019 has been

produced through a neural network trained against SMAP radiometer-based descending

SSM (hereafter referred to as "NN-SM product") (Yao et al., 2021). The global

validation results show that this NN-SM product is better than the JAXA and LPRM

products.

Besides, the NN-SM has also been compared with another long-term ~25-km all-

weather SSM dataset generated through the European Space Agency (ESA)'s Climate

Change Initiative (CCI) program. The ESA-CCI SSM product is different from the rest

products mentioned above in that it was implemented by fusion of observations from

comprehensive AM- and PM-based satellite sensors, rather than only relying on the

radiometers of AMSR series. According to Yao et al. (2021), the ESA-CCI SSM has

173   slightly better validation accuracy than the NN-SM product, but the number of available

174   observations per pixel cell in an entire year is much smaller for the ESA-CCI SSM in

175   Southeast China. In view of all above coarse-resolution SSM data products, we finally

176   selected the NN-SM product to implement the following spatial downscaling

177   procedures rather than the ESA-CCI SSM, to make a balance between data accuracy

178   and data availability per year. We have also made additional evaluations within China

179   in Section Appendix-A to ensure the relatively outstanding performance of the NN-SM

180   product as described above.

181   Table 1 Information of all-weather microwave remote sensing coarse-resolution SSM data

182   products that can be potentially downscaled to obtain fine resolution SSM.

Name	Resolution	Satellite radiometers involved	Data availability (URL)
NN-SM product	36 km (by the EASE Grid projection)	AMSR-E/ AMSR-2 (2002-2011, 2012-present)	https://data.tpdc.ac.cn/en/data/c26201fc-526c-465d-bae7-5f02fa49d738/
ESA-CCI v6.1 product	0.25°	AMSR-E/ AMSR-2/ SMOS/ WindSat/ SMMR/ SSM/I/ TMI (1978-2020)	https://www.esa-soilmoisture-cci.org/v06.1_release
JAXA product	0.25° / 0.1°	AMSR-E/ AMSR-2 (2002-2011, 2012-present)	https://gportal.jaxa.jp/
LPRM product	0.25° / 0.1°	AMSR-E/ AMSR-2 (2002-2011, 2012-present)	https://search.earthdata.nasa.gov/
UMT product	25 km (by the EASE Grid projection)	AMSR-E/ AMSR-2 (2002-present)	http://files.ntsg.umt.edu/data/LPDR_v2/

183

184   2.1.2 Optical remote sensing data and digital elevation model (DEM)

185   Optical remote sensing datasets provide finer spatial texture information on the

186   daily basis for the downscaling purpose of PM SSM. Such data that can be used as

187   inputs of our SSM product processing line are mainly provided by the Moderate-

188   resolution Imaging Spectroradiometer (MODIS) onboard the Terra and Aqua satellites.

189   Specifically, they involve the 1-km daily night-time Aqua MODIS LST product

190   (MYD21A1N.v061) and the 500-m daily "Bidirectional Reflectance Distribution

191   Function (BRDF)" - Adjusted Reflectance dataset (MCD43A4.v061). MYD21A1 LST

192   data can be recognized as a crucial proxy of land surface thermal capacity (Fang et al.,

193   2013) and soil evaporative rate (Merlin et al., 2008). The MCD43A4 nadir reflectance

194   product, with view angle effect corrected using the BRDF model, is capable to provide

195   observations from visible to shortwave-infrared bands that can characterize water

196   content variation of the bare soils as well as the vegetation canopy. Overall, the above-

197   mentioned datasets were selected primarily because they deliver indicators (land

198   surface thermal capacity, soil evaporative rate, or vegetation condition) that can well

199   response to soil moisture dynamics from different aspects. Prior to being employed for

200   SSM downscaling, conventional pre-processing procedure of pixel quality check was

201   applied for both optical products by screening out pixels not classed as "good quality",

202   according to the 8-bit "Quality Assessment (QA)" field of each spectral band. Moreover,

203   to normalize their natively different spatial resolutions, all MCD43A4 based reflectance

204   values at the 500-m pixel level were upscaled to the sinusoidally projected MODIS 1-

205   km grids using their spatial averages.

206   Apart from MODIS optical remote sensing data, all 90-m DEM tiles generated by

207   the NASA Shuttle Radar Topography Mission (SRTM; http://srtm.csi.cgiar.org/, last

208   access: July 10, 2020) were mosaicked all over China and then employed as another

209   essential input variable for the procedures as described by Section 2.2.2 below. Similar

210   to that applied to the MCD43A4 product, spatial upscaling in correspondence to the

211   MODIS 1-km grids is also an indispensable pre-processing step for the mosaicked

212   DEM data.

213   2.1.3 Study area and Ground validation data

214   Our study area is set up as the total terrestrial extent of China. To comprehensively

215   evaluate the SSM downscaling performances for different geographic settingsregions

216   (see Section 3.3), we divided the country further into six different geographic-climate

217   regions based on various features involvingusing elevation, rainfallprecipitation,

218   hydrogeology, vegetation type, and topography. Particularly, theyThe six regions

219   include the Northeast Monsoon (NEM) region, the Northwest Arid (NWA) region, the

220   Qinghai–Tibet Plateau (QTP) region, the North China Monsoon (NCM) region, the

221   South China Monsoon (SCM) region, and the Southwest Humid (SWH) region. The

222   detailed delimitation principle of these geographic-climate regions was originally

223   described in Meng et al. (2021). The geographic zoning map is shown in Fig. 1, while

224   the corresponding shapefile boundary files can be derivedaccessed from the Resource

225   and Environment Science and Data Center of the Chinese Academy of Sciences

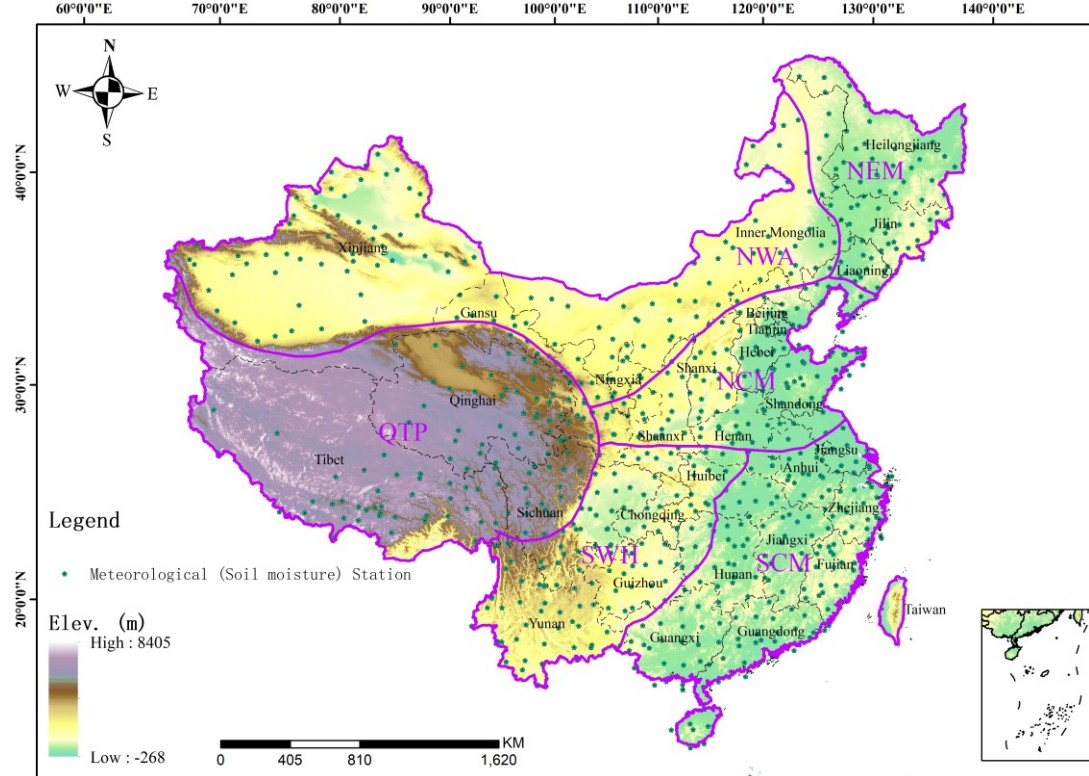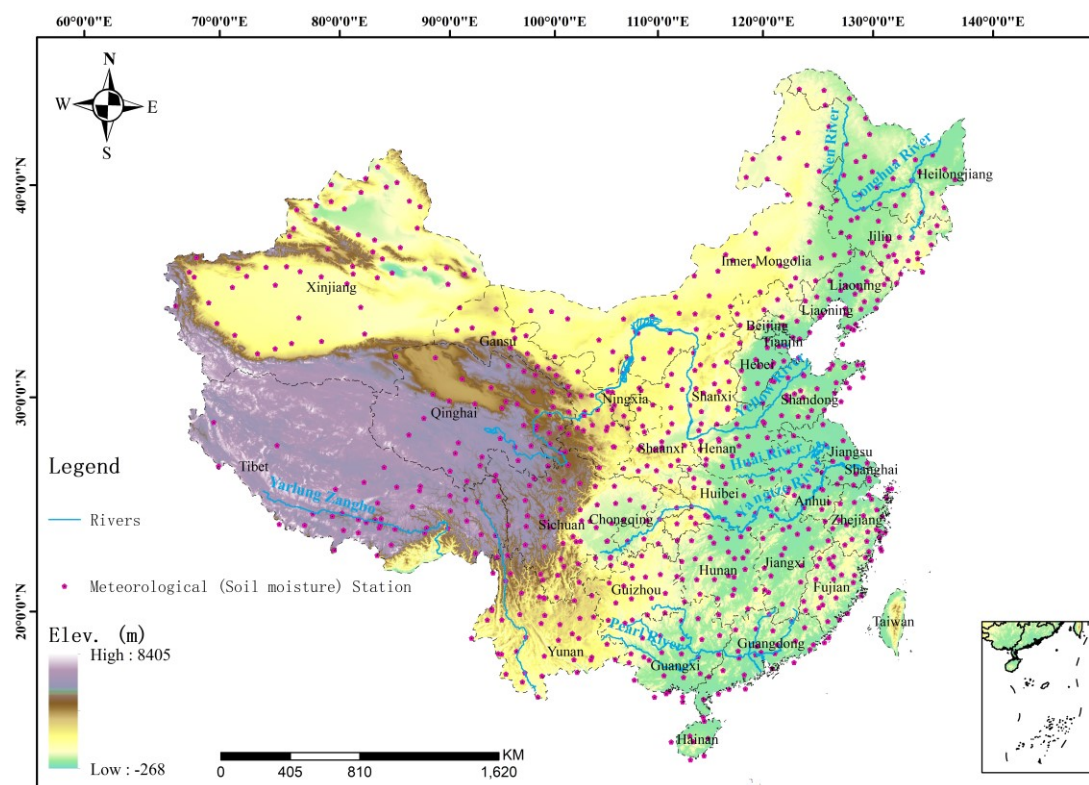226   (http://www.resdc.cn/, last access: May 22, 2021).

227

Fig. 1 The ~~provincial-level administration~~geographic zoning map of China (delineated using the purple

color) superposed with topographic information, as well as general locations for the 756 basic

232   meteorological stations (http://data.cma.cn/, last access: January 20, 2021) that provide partial

233   benchmark measurements for SSM and LST validation in this study.

234   We utilized ground soil moisture measurements for validating the downscaled

235   remote sensing SSM product at the local scale. The ground measurements are derived

236   from 2417 meteorological stations (including 756 basic stations of the National Climate

237   Observatory and 1661 regionally intensified stations) of over China, as partially shown

238   in Fig. 1Fig. 1. The soil moisture measurement devices in these stations, with uniform

239   observation standards, are instrumented under the national project of "Operation

240   Monitoring System of Automatic Soil Moisture Observation Network in China (Wu et

241   al., 2014   )", the construction of which has been led by China Meteorological

242   Administration since 2005. Until 2016, all stations have been in operation for

243   automatically observing hourly in situ soil moisture dynamics at eight different depth

244   ranges (0-10 cm, 10-20 cm, 20-30 cm, 30-40 cm, 40-50 cm, 50-60 cm, 70-80 cm, 90-

245   100 cm). It has also been widely used by previous studies for evaluating satellite soil

246   moisture estimates in China (Meng et al., 2021; Chen et al., 2020; Zhang et al., 2014;

247   Zhu and Shi, 2014). In our current study, ground measurements matching the shallowest

248   depth range (0-10 cm) from the initial time of each station until the end of 2019 are

249   employed as validation benchmark of the satellite SSM retrievals. At the temporal

250   dimension, measurements made at 1:00 A.M. and 2:00 A.M are averaged, in order to

251   match the mean satellite transit time of 1:30 A.M. for AMSR descending observations.

252   Moreover, 0-cm top ground temperatures are simultaneously measured at all these

253   meteorological stations on the daily basis, at the local time windows of 2:00 A.M./P.M.

and 10:00 A.M./P.M., respectively. We therefore exploited such measurements recorded at 2:00 A.M. to validate the cloud gap-filled night-time (~1:30 A.M.) LST estimates over the Aqua-MODIS based 1-km pixels containing these stations (see Section 2.2.2). Our primary validation period covers the entire years of 2017, 2018, and 2019.

In addition to the ground soil moisture measurements, the SMAP Level3 radiometer-based daily 36-km SSM product (https://dx.doi.org/10.5067/OMHVSRGFX38O) in its descending orbit scenes (at ~6:00 A.M. of local time) from 2016 to 2019, was employed as another complemental validation benchmark. This dataset is potential for providing more comprehensive evaluations to our developed product at regional/national scales, especially on account of its notably creditable performance (see Fig. A1 in Appendix A). The latest version of this dataset (SPL3SMP, Version 8) contains soil moisture retrievals based on different algorithms including the dual channel algorithm and the single channel algorithm. In this study we only extracted SSM estimates derived with the dual channel algorithm because this algorithm iwas reported to slightly surpassoutperform the single channel algorithm over some agricultural cropland core validation sites (O'neill et al., 2021).

2.1.4 Ancillary SSM products for comparison

In order to comprehensively demonstrate the validation performance of our proposed SSM product, there is necessity to make an inter-comparison against similar

275 existing datasets. In this regard, we introduced the Level2 SMAP/Sentinel Active-

276 Passive combined SSM product on 1-km earth-fixed grids, i.e., the SPL2SMAP_S_V3

277 dataset (Das et al., 2020), and used its validation performance against in-situ

278 measurements throughout the years of 2017, 2018, and 2019, as a baseline to better

279 evaluate our proposed SSM product. The SPL2SMAP_S_V3 dataset contains global

280 SSM at resolutions of 3 km and 1 km respectively, which were disaggregated from the

281 SMAP radiometer-based SSM retrievals of 36-km/9-km footprints in conjunction with

282 the high-resolution Sentinel-1 C-band radar backscatter coefficients (Das et al., 2019).

283 To our knowledge, this dataset is possibly the only publicly available product which

284 can provide global remote sensing SSM estimates at the 1-km resolution. The sentinel

285 backscatter coefficient inputs for this product are only those received in the descending

286 orbit scenes (at ~6:00 A.M. of local time), whilst the closest SMAP SSM retrievals

287 from either ascending (at ~6:00 P.M. of local time) or descending orbits are used to

288 spatially match up with the sentinel-1 scene. It is noticed that at the descending

289 observation time the soil moisture vertical profile has approached a hydrostatic balance

290 (Montaldo et al., 2001), thereby providing the optimal chance for soil moisture fusion

291 and validation with observations at different soil depths. Therefore, we only selected

292 the 1-km disaggregated SSM estimates based on descending SMAP SSM retrievals (i.e.,

293 the subset with field name of 'disagg_soil_moisture_1 km' in the SPL2SMAP_S_V3

294 dataset). Meanwhile, the 0-10 cm in-situ soil moisture measurements observed at 6:00

295 A.M. and the SMAP radiometer-based descending SSM estimates were employed as

296 the validation benchmarks, in a manner similar to that applied to our proposed SSM

297 product (Section 2.1.3).

298 ~~(O'neill et al., 2021)~~

2.2 Methodology

300 The general methodological framework for producing the all-weather daily 1-km

301 SSM product is shown as in Fig. 2~~Fig. 2~~, with details described in the following context
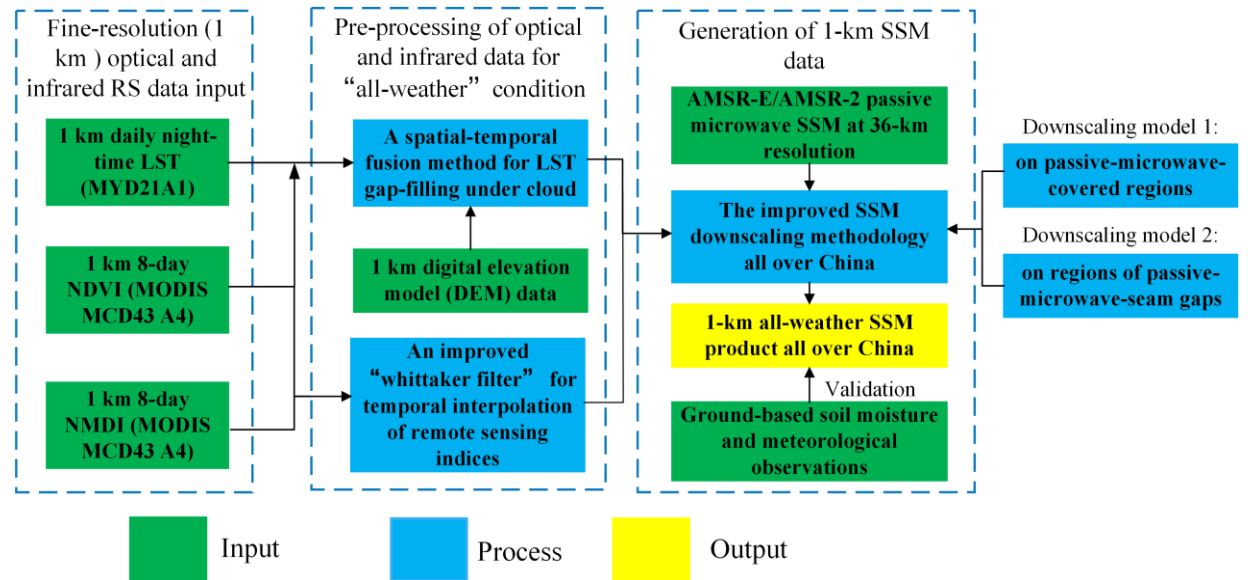
302 of this section.

303
304 Fig. 2 The overall methodological framework of this study.

305 2.2.1 Reconstruction of thermal-infrared LST and remote sensing (vegetation)

306 indices under cloud

307 Reconstruction of missing pixels under cloud in the optical remote sensing input

308 datasets is the prerequisite for achieving the "all-weather" property of the final

309 downscaled SSM output. For reconstructing thermal-infrared LST, we adopted the

310 cloud gap-filling method as proposed by our previous study (Song et al., 2019a). This

311 method, also referred to as a typical "spatio-temporal data fusion" (STDF) method

312 (Dowling et al., 2021), was built using clear-sky LST observations of spatially

313 neighboring pixels observed at proximal dates, with concurrent NDVI and DEM also

314 employed as additional data inputs. The STDF method can be expressed as follows:

315 $$LST^*_{t_1} = a \times LST^*_{t_0} + b \times NDVI^*_{t_1} + c \times DEM^* + d \qquad (1)$$

316 Where the superscript "*" indicates that this variable has been normalized to the range

317 0 to 1.0 (Song et al., 2019a), based on the maximum and minimum values of that

318 variable found across China (excluding invalid values representing states of snow, ice,

319 and water bodies). Parameters $a, b, c,$ and d are coefficients fitted between all pixels

320 with clear-sky LST estimates on a specific date t_1 ($LST^*_{t_1}$) and their counterparts on

321 one proximal date, t_0 ($LST^*_{t_0}$). $NDVI^*_{t_1}$ indicates the corresponding (normalized) NDVI

322 on the t_1 date calculated using the MCD43A4 daily product. After deriving the

323 coefficients of $a, b, c,$ and d, Equation (1) was used to fill all cloudy MODIS LST pixels

324 on the t_1 date. For any t_1 date included in the study period, the t_0 date was iterated among

325 all neighboring dates of t_1 meeting the condition | t0- t1|<=30 (from the nearest date to

326 the furthest date). The average of estimated LST values for t_0 was then taken where a

327 cloud gap pixel was filled more than once (based on the iterative t_0 dates). The iteration

328 was stopped when the fraction of pixels with effective LST values on t_1 was equal to or

329 exceeded 0.99.

330 An important flaw of this STDF method should be noticed with regard to

331 potentially existential bias of the cloud gap-filled LST outputs, because the outputs

332 represent theoretically reconstructed LST under clear sky rather than under the real

333 cloudy condition. Another of our previous studies (Dowling et al., 2021) concerning

334 this STDF method proposed a follow-up step, which incorporated PM-derived surface

335 temperature, to adjust that bias. In our current production pipeline, however, this

336 follow-up step for cloud bias adjustment in LST was not carried out. This is because

337 the results in Section Appendix-B show that using LST generated by the STDF alone

338 leads to more accurate SSM outcomes in general. The possible reasons for this are

339 discussed in Section 4.2.

340 Reconstruction of the remote sensing vegetation indices under cloudy conditions,

341 including NDVI and MNDI, was simply based on the modified time series filter of the

342 Whiitaker Smoother (MWS) as developed by Kong et al. (2019). This is reasonable

343 because the dynamic trends of vegetation growth are relatively less volatile compared

344 to LST on the daily basis, and can thus be gap-filled for missing values using a time-

345 series-filtering-like algorithm.

346 2.2.2 Improved downscaling technique of SSM based on fusion of PM and

347 optical/infrared data

348 The core component of the SSM downscaling methodology is an improved linking

349 model between PM SSM and (fine-resolution) optical remote sensing observations.

350 This model enhances the relatively poorer performance of the conventional DISPATCH

351 in energy-limited regions, whilst maintains the generally good quality of the

352 DISPATCH in water-limited ones. Therefore, the improved model is more appropriate

353 to be applied in China which contains a wide range of geographical settings, compared

354 to other conventional downscaling models. Since this model origins from our previous

355 study (Song et al., 2021), herein we simply give its mathematical expression as follows:

356
$$SSM = \frac{a \times \ln(1 - SEE)}{1 - b \times NMDI} + c \qquad (2)$$

357 In Equation (2), *SEE* denotes "soil evaporative efficiency" and is a mathematical

358 function of LST and the typical Normalized Difference Vegetation Index (NDVI), with

359 its specific form described in Merlin et al. (2008). NMDI is another remote sensing

index calculated as $\frac{R_{infr,860nm}-(R_{sw,1600nm}-R_{sw,2100nm})}{R_{infr,860nm}+(R_{sw,1600nm}-R_{sw,2100nm})}$ (Wang and Qu, 2007).

$R_{infr,860nm}$, $R_{infr,1600nm}$, and $R_{infr,2100nm}$ represent land surface reflectance signals derived from three different MODIS-MCD43A4 based near-infrared/shortwave-infrared bands, with their wavelengths centering at 860 nm, 1600 nm, and 2100 nm respectively. The parameters *a, b,* and *c* are empirical coefficients that represent background information of local soil texture and vegetation types. In Song et al. (2021), these coefficients have been fitted and calibrated based on multi-temporal observations at the PM pixel scale. In our current study, however, we have discovered that coupling of multiphase observations at both the spatial and the temporal dimensions can lead to more optimal solution of the coefficients, as they can produce downscaled SSM images with notably declined effect of 'mosaic' against the original PM 36-km pixels. Therefore, the modified optimal cost function χ^2 for deriving these coefficients is re-defined as follows:

$$\chi^2 = \sum_{d=-dl}^{dl} \sum_{i=0}^{N=ws \times ws} w_i \times (SSM_{ob,i,d} - SSM_{mod,i,d})^2 \qquad (3)$$

Through the cost function, the spatial extent of each 36-km pixel P_0 on any arbitrary date D_0 obtains a unique set of coefficients. As shown by Equation (3), all pixels were exploited within the spatial square window (with its side length equal to *ws*) centered at P_0 ranging from *-dl*-th day to *dl-th* day relative to the date of D_0. To determine the optimum values for *dl* and *ws*, we have tested each member in the collection of [3, 5, 7, 9, 11, 13] for both of the parameters. Evaluation against in-situ data indicates that the optimum *dl* and *ws* are 5 and 7, respectively (results are similar to what is shown in Section 3.2, but not presented here). *SSM_{ob}* and *SSM_{mod}* denote the AMSR NN-SM 36-

382 km SSM observations as well as SSM observations modelled by Equation (2) based on

383 upscaled optical datasets, respectively. w_i is a weight coefficient used to ensure that

384 neighboring observations near the centering pixel P_0 play more dominating roles as

385 compared with the far-end pixels in the cost function, considering the "Tobler's First

386 Law of Geography (Sui, 2004)" . w_i is calculated using an adaptive bi-square function:

387
$$w_i = [1-(\frac{dis_i}{b})^2]^2, dis_i < b$$
$$w_i = 0, dis_i >= b$$
(4)

388 where dis_i indicates the distance between the i-*th* pixel and the centering pixel P_0. *b* is

389 named as the adaptive kernel bandwidth of the bi-square function (Duan and Li, 2016),

390 and is optimized as 200 km through using a cross validation method as recommended

391 by Brunsdon et al. (1996).

392   With the linking model obtained, we can subsequently utilize the spatial

393 downscaling relationship function to produce 1-km fine resolution SSM. The

394 downscaling relationship function is constructed by transforming the linking model into

395 its Taylor expansion formula and preserving all components with respect to the input

396 optical variables of the linking model at first and second orders. This relationship is

397 inspired from Malbéteau et al. (2016) and Merlin et al. (2010), and is mathematically

398 described below:

399 $SSM_{1\text{-}km}=SSM_{36km}+(\frac{\partial SSM}{\partial SEE})_{36km} \times (SSE_{1km} -< SSE >_{36km}) + 0.5 \times \left(\frac{\partial^2 SSM}{\partial SEE^2}\right) \times (SSE_{1km} -$

400 $< SSE >_{36km})^2 +(\frac{\partial SSM}{\partial NMDI})_{36km} \times (NMDI_{1km} -< NMDI >_{36km}) + 0.5 \times \left(\frac{\partial^2 SSM}{\partial NMDI^2}\right) \times$

401 $(NMDI_{1km} -< NMDI >_{36km})^2$ (5)

402 In the above relationship, $<>$ denotes the spatial averaging operator for all of the 1-km

403 optical remote sensing input variables within the corresponding 36-km pixel,

404 $\frac{\partial SSM}{\partial SEE}\left(\frac{\partial^2 SSM}{\partial SEE^2}\right)$ and $\frac{\partial SSM}{\partial NMDI}\left(\frac{\partial^2 SSM}{\partial NMDI^2}\right)$ respectively denoting the first-(second-) order

405 partial derivative of the linking model described in Equation (2).

406   It should be noticed that there exist middle-/low-latitude gap regions between

407 seams of neighboring daily AMSR-E(-2) swaths, indicating that SSM_{36km} in Equation

408 (5) is not always available on the daily basis (Song and Zhang, 2021a). For such PM-

409 seam gaps on a particular date t_0, the corresponding $SSM_{36km,t0}$ in Equation (5) is

410 substituted by $0.5 \times (SSM_{36km,t0+1} + SSM_{36km,t0-1}) + \Delta SSM_{36km,t0}$. Herein $SSM_{36km,t0-1}$

411 and $SSM_{36km,t0+1}$ respectively denote the SSM estimate before and after the date of t_0.

412 $\Delta SSM_{36km,t0}$ is a component for correcting inter-day bias, with the following expression:

413
$$
\begin{aligned}
\Delta SSM_{36km,t0} = {}& SSM\left(SEE_{36km,t0},\ NMDI_{36km,t0}\right) - \\
& 0.5 \times \left(SSM\left(SEE_{36km,t0-1},\ NMDI_{36km,t0-1}\right) + SSM\left(SEE_{36km,t0+1},\ NMDI_{36km,t0+1}\right)\right)
\end{aligned}
\tag{6}
$$

414   In the above equation, $SSM(SEE_{36km},\ NMDI_{36km})$ denotes SSM that is directly

415 modelled based on Equation (1) using 36-km SEE and NMDI. The 36-km SEE and

416 NMDI are obtained via averaging the variables spatially from their native resolution at

417 1-km. If all $SSM_{36\text{-}km}$ during the three consecutive days (t_0-1, t_0, and t_0+1) are missing

418 due to other extreme conditions like snow, ice, or surface dominated by substantially

419 large water bodies, the downscaling process cannot be fulfilled and all 1-km sub-pixels

420 with the $SSM_{36\text{-}km}$ have to be set as null values.

421 2.2.3   Evaluation metrics

422   We employed the classic metrics of 'Root Mean Square Difference (RMSD)' and

423 correlation coefficient (r-value) for evaluating satellite-based (SSM and LST) estimates

424 against ground measurements. Herein RMSD is not referred to as 'Root Mean Square

425 Error (RMSE)', although the latter term shares the same definition and has been used

426 more commonly in previous studies. This is because both ~~the~~ ground observations

427 ~~benchmark data~~ and other benchmark data (i.e. SMAP radiometer-based SSM) may

428 also present measurement uncertainties in practice. For SSM evaluation, the unbiased

429 RMSD, or ubRMSD (Entekhabi et al., 2010a; Molero et al., 2016), is calculated instead

430 of RMSD when validated against ground soil moisture measurements. This can~~in order~~

431 ~~to~~ better investigate the time series similarity between satellite and ~~ground soil~~

432 ~~moisture~~in situ datasets by eliminating the systematic bias caused by spatial scale

433 mismatch between them.

434   The above-mentioned classic metrics are primarily suitable to evaluate the

435 absolute reliability of an independent remote sensing product. However, we also require

436 another metric for characterizing the relative improvement of the downscaled SSM

437 estimates against the original PM observations on capturing local soil moisture

438 dynamics. For this purpose, we employed the "gain metric" of G_{down}, which was

439 developed particularly by Merlin et al. (2015) for assessment of soil moisture

440 downscaling methodology. G_{down} is a comprehensive indicator for evaluating gains of

441 the downscaled SSM against the original coarse-resolution PM data in terms of their

442 mean bias, bias in variance (slope), and time series correlation with ground benchmark.

443 It has a valid domain between -1 and 1, with positive (negative) value indicating

444 improved (deteriorated) spatial representativeness of the downscaled SSM against the

445 original PM data. Detailed definition and introduction of G_{down} are given in Equation

446 (8) and Section 3.3 of Merlin et al. (2015).

3. Results

3.1 Evaluation on reconstructed thermal-infrared LST under cloud

450   The meteorological-station-based validation of reconstructed 1-km thermal-

451 infrared LST under cloud were preliminarily fulfilled, to ensure the high quality of input

452 dataset variables for SSM downscaling. Since disadvantageous effects might be

453 brought to this validation campaign by the potentially existing heterogeneity of the

454 validated 1-km thermal-infrared remote sensing pixels, we firstly analyzed correlations

455 between estimated and benchmark datasets at each station, only based on satellite

456 remote sensing observations obtained under clear sky. Stations that have their

457 correlation coefficients (r_{clr}) lower than 0.9 herein have to be screened out because there

458 exist higher chances of cross-scale spatial mismatch within and around these stations

459 in terms of the land surface thermal properties. Among all 2417 stations (see Section

460 2.1.3) where 0-cm in-situ top-ground temperature measurements were available, we

461 finally preserved 2107 stations characterized by r_{clr} >0.9. In the subsequent step, remote

462 sensing LST under cloud and under clear-sky conditions were respectively validated at

463 these stations, with the results revealed in Fig. 3Fig. 3. It is manifested through Fig.

464 3Fig. 3-(a) and -(b) that very close performances have been achieved between the clear-

465 sky and the cloudy scenarios, especially considering their almost equally high

23

466   validating correlations between 0.94-0.96. For each independent station, we calculated

467   the "RMSD difference (RMSD_diff)" between the two scenarios, based on the formula

468   of "$RMSD_{clr}$- $RMSD_{cld}$ (the subscripts of *'clr' and 'cld'* denote clear-sky and cloudy

469   conditions separately)". The statistical distribution of this RMSD difference with regard

470   to different stations is shown in Fig. 3Fig. 3-(c). Apparently, 1942 stations all over the

471   country have obtained an RMSD difference value below 2.6 K, and the mean RMSD

472   difference is about 1.9 K. All above results have indicated that the uncertainty of our

473   night-time LST reconstruction algorithm proposed for cloudy conditions is not very

474   significant. The corresponsive uncertainty that could be propagated to downscaled SSM

475   in this stage is analyzed below in Section 3.2.

(a) Clear-sky LST validation

(b) Cloudy-sky LST validation

(c) Statistical distribution of station-based LST validation results
(2107 stations qualified for $r\_clr>0.9$)

Range of RMSD_diff	Station numbers
<0 K	66
0 -1.6 K	367
1.6-2.6 K	1575
2.6-3.7 K	99
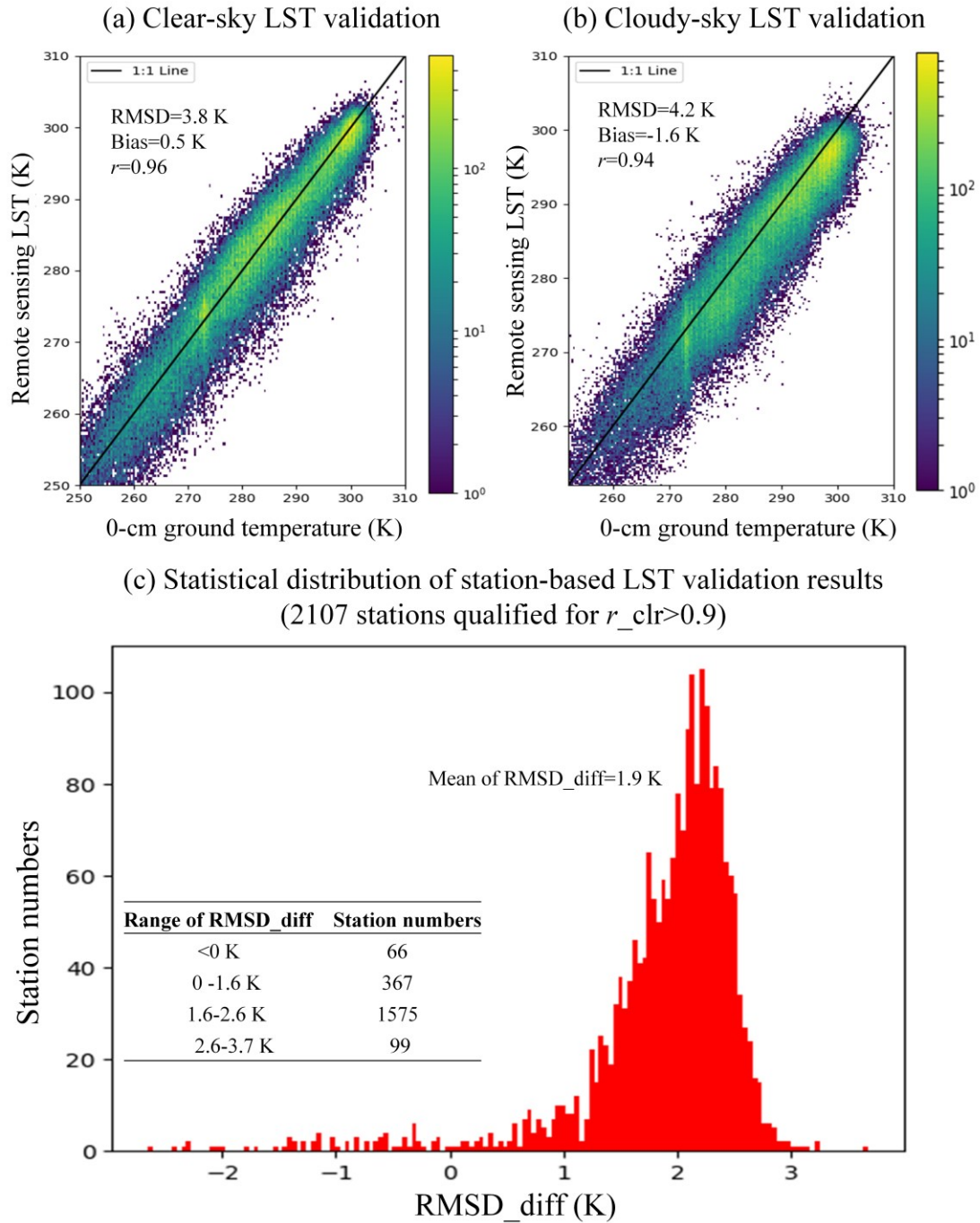
Fig. 3 validation results of the cloud gap-filled LST in China. (a) Density plot of thermal infrared LST under clear-sky condition compared to the 0-cm ground temperature measurements for all stations. (b) Same to (a) but for thermal infrared LST under cloudy conditions. (c) Statistical distribution of difference between RMSD of clear-sky LST and RMSD of gap-filled LST under cloudy condition with regard to different meteorological stations over the study region.

3.2 Evaluation on the final 1-km SSM product

The overall validation results of the finally downscaled 1-km SSM product against ground soil moisture data is demonstrated in Fig. 4~~Fig. 4~~. Fig. 4~~Fig. 4~~-(a) shows that about 85% (N: 1833) of the total 2154 stations (the remaining 263 stations are located in pixels with no effective PM observations and are thus removed) have obtained significantly positive downscaling gains (G_{down}>0.03). This hints that the 1-km SSM product can better capture the dynamic behaviors of local ground soil moisture data than the original 36-km PM NN-SM data, revealing higher spatial representativeness of the downscaled SSM data product over the country. According to Fig. 4~~Fig. 4~~-(b), the mean ubRMSD of all stations is about 0.054 vol/vol, while 90% of those stations have the number lower than 0.088 vol/vol. In addition, we made another analysis concerning the possible influence of land cover types on SSM downscaling performance in Fig. 4~~Fig. 4~~-(c). The spatial information of land cover types was derived from the MODIS MCD12Q1 (10.5067/MODIS/MCD12Q1.006) IGBP-based land use image in 2019. For stations that experienced land use change throughout the years of the study period, the ubRMSD is only reported for data in the year of 2019. Clearly, better accuracies are observed mainly in grassland, cropland and bare soil surface, whilst relatively poorer performances (with averages of ubRMSD higher than 0.06 vol/vol) are seen in urban regions, (woody) savanna, and crop-to-natural-vegetation mosaic areas. Such a relative performance across land covers is logical because all the land cover types with their average ubRMSD higher than 0.06 vol/vol are characterized

504   by lower hydrologic homogeneity in terms of their definition, e.g. savanna, which is a

505   mixture of grass and tall trees, and urban areas, which are composed of impervious
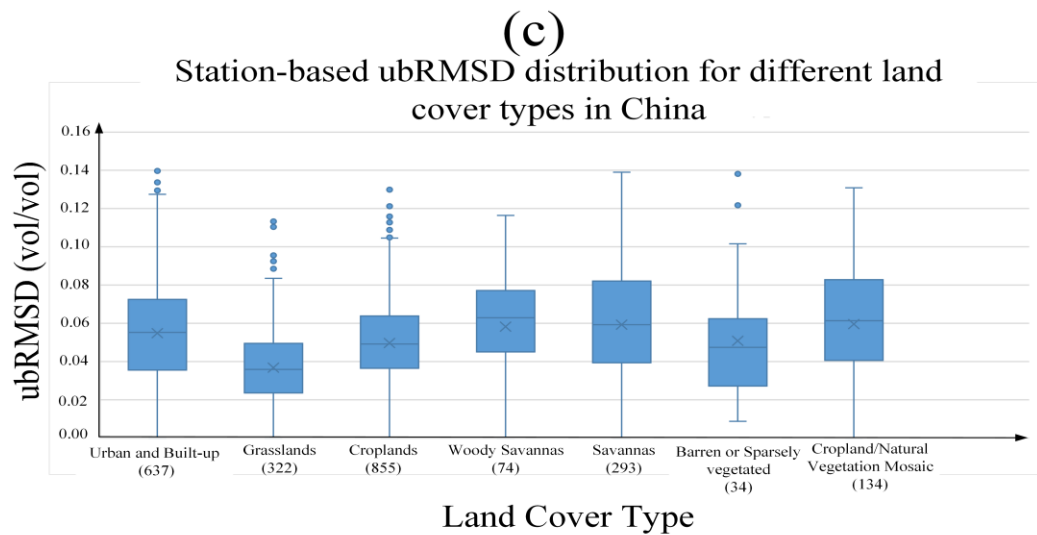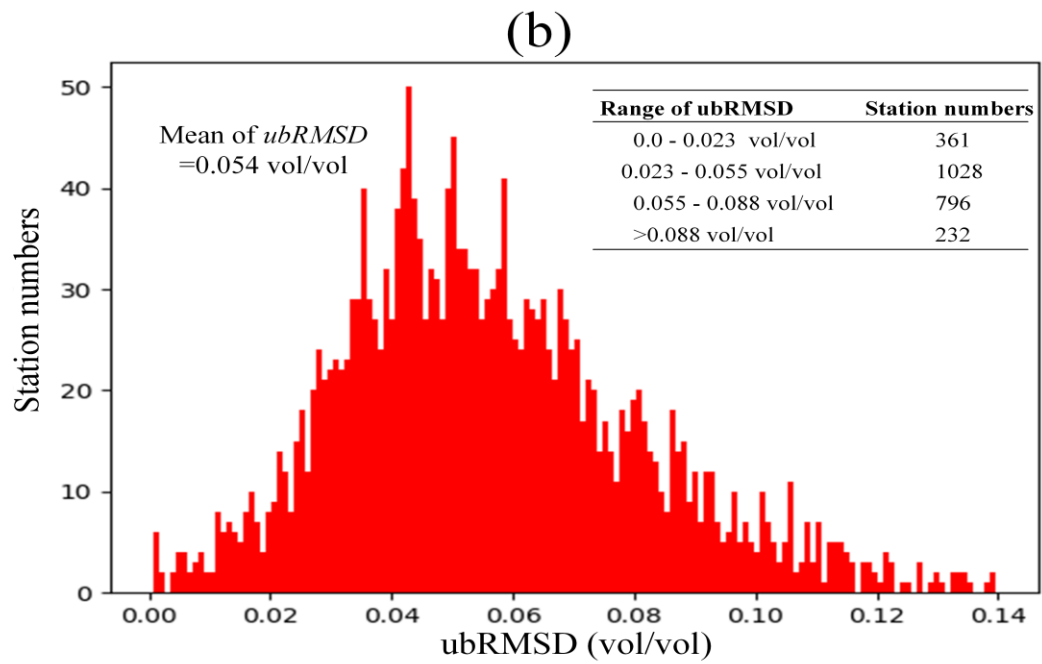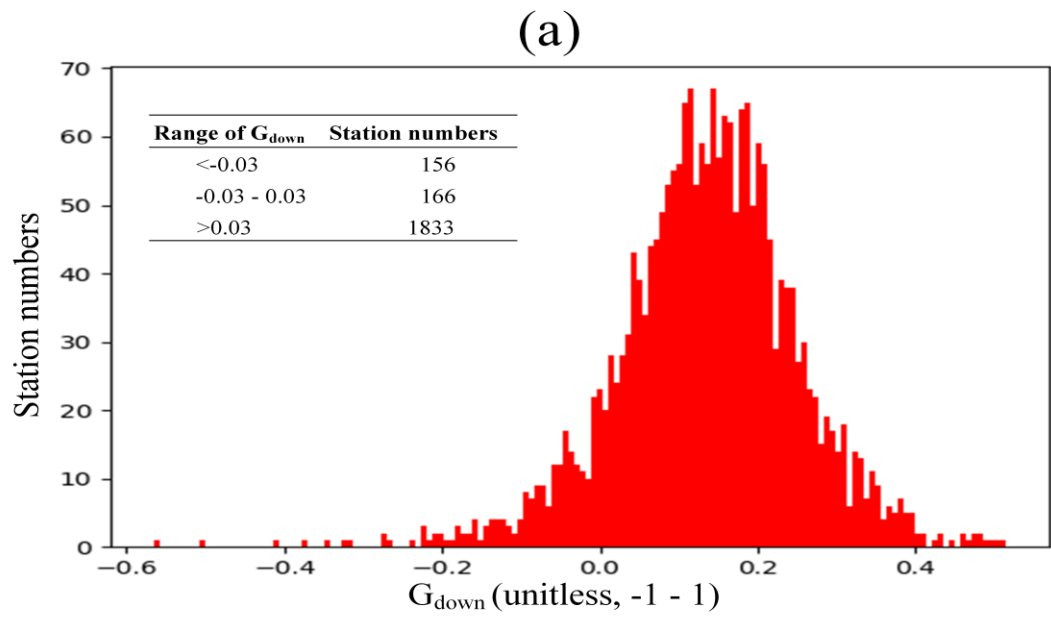
506   underlying surface.

(a)

Range of G_{down}	Station numbers
<-0.03	156
-0.03 - 0.03	166
>0.03	1833

G_{down} (unitless, -1 - 1)

Station numbers

(b)

Mean of *ubRMSD* =0.054 vol/vol

Range of ubRMSD	Station numbers
0.0 - 0.023 vol/vol	361
0.023 - 0.055 vol/vol	1028
0.055 - 0.088 vol/vol	796
>0.088 vol/vol	232

ubRMSD (vol/vol)

Station numbers

(c)

Station-based ubRMSD distribution for different land cover types in China

ubRMSD (vol/vol)

Land Cover Type

508   Fig. 4 General validation results of the currently developed SSM product. (a) G_{down} distribution for

509   different stations over China. (b) ubRMSD distribution for different stations over China. (c) ubRMSD

510   statistics reported for different land covers. The numbers in the parentheses of the x-axis labels

511   represent the amount of meteorological stations corresponding to that specific land cover type.

512   In Fig. 5, we ~~demonstrated~~compared time series ~~features~~ of regionally aggregated

513   SSM from our developed 1-km SSM product ~~as well as from~~to that from the SMAP

514   36-km descending SSM, for each of the six different geographic-climate regions (as

515   shown in Fig. 1) from 2016 to 2019. Via this effort, we mainly aim to reveal the

516   consistency degree on reflecting soil moisture temporal dynamics at different

517   geographical settings between the two SSM products. This also provides another view

518   to evaluate the reliability of our developed product. Because the SMAP radiometer has

519   a slightly longer revisit cycle (~2-3 days) than AMSR-2, the time series data are also

520   aggregated and averaged at the temporal dimension, with a displayed revisit cycle equal

521   to three days. Overall, the time series data correlate well with each other for all six

522   regions. The relatively lower RMSDs (<0.02 vol/vol) are found in regions with

523   comparatively sparser vegetation covers including the NWA region, the QTP region,

524   and the NCM region. For other three dense-vegetation regions, the performances of our

525   developed product are slightly ~~worse~~poorer. This is especially the case for the SCM

526   region, with a lower r-value of 0.84. The reason can be attributed to the enlarged

527   difference on penetration depth into the soil layers between L-band (SMAP) and C-/X-

528   /K- band (AMSR-2) emissions under dense vegetation covers (Ulaby and Wilson,

529   1985).

(a) in NEM
RMSD=0.024 vol/vol
r=0.90

(b) in NWA
RMSD=0.008 vol/vol
r=0.95

(c) in NCM
RMSD=0.015 vol/vol
r=0.94

(d) in QTP
RMSD=0.017 vol/vol
r=0.95

(e) in SWH
RMSD=0.021 vol/vol
r=0.90

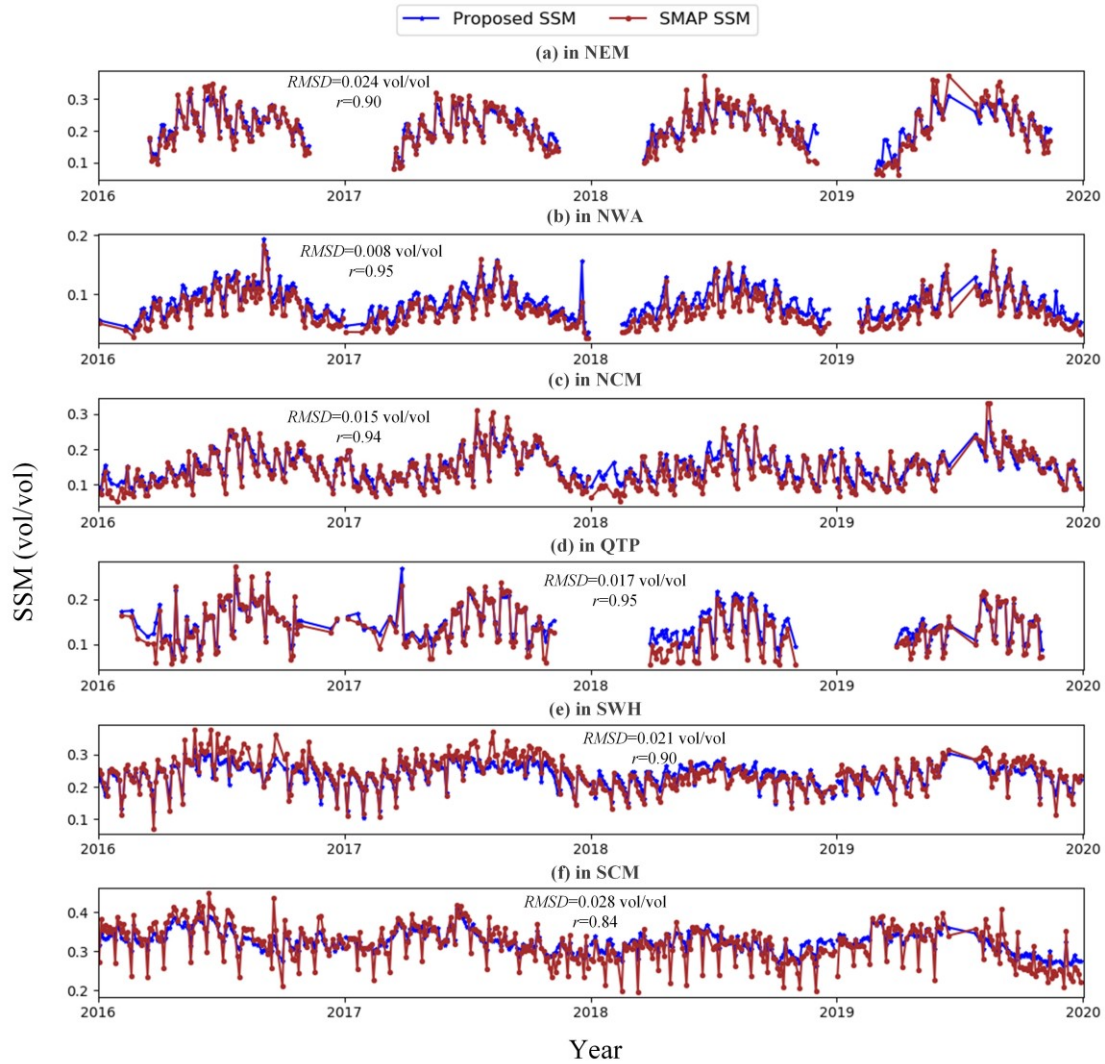**(f) in SCM**
RMSD=0.028 vol/vol
r=0.84

Fig. 5 Time series of SSM aggregated at each of the 6 different geographic-climate regions (as shown in Fig.1) in China for our developed 1-km product as well as for the SMAP 36-km SSM dataset. The time series range from 2016 to 2019, with a revisit cycle of three days.

In Fig. 6-(b) we employed the downscaled SSM image on ~~April~~ May 29, 2018, as an example to demonstrate the spatial features of the developed product. Meanwhile, we also show the map of SMAP/Sentinel combined SSM (SPL2SMAP_S_V3) obtained from ~~May~~April 26 to ~~May~~April ~~31~~1, 2018 in Fig. 6~~Fig. 5~~-(a), as a contemporaneous comparison reference. Clearly, the SPL2SMAP_S_V3 map has a much lower coverage percentage over the study region compared with the map of the currently developed

540	product on one single date, even though the former was generated based on multi-date

541	images. Both maps show similar spatial texture depicting the relatively dry climate in

542	northwestern China compared with the humid climate in the Middle-lower Yangtze

543	River Plain. Nevertheless, there also exist cases where the details in texture differ

544	prominently, like that in the far northeastern end of the country.

545	For the sake of further analysis on this point, results of the quantitative comparison

546	as proposed in Section 2.1.3 and Section 2.1.44, is demonstrated in Fig. 6Fig. 5-(c),

547	and Fig. 5-(d), -(e), and -(f). Fig. 6-(c) and -(d) show the RMSD maps of the two

548	respective products against SMAP radiometer-based SSM estimates at the 36-km pixel

549	scale. For both products it is manifested that compared with the lower averaged RMSD

550	of 0.04 vol/vol in the NWA region, the uncertainty can increase (shown in yellow) in

551	the densely vegetated NEM and the SCM regions, with averaged RMSDs of 0.07-0.08

552	vol/vol. However, our developed product has significantlynoticeably lower RMSD

553	(0.05 vol/vol) than the SPL2SMAP_S_V3 data (0.07-0.09 vol/vol) in the SWH and part

554	of the QTP regions. Considering their relatively higher elevations, it may be roughly

555	drawn that our downscaled SSM product is more reliable than that downscaled based

556	on active-passive microwave combined datasets in areas with increased topographic

557	effects. Fig. 6-(e) shows that tThe currently developed SSM product obtained a 0.078

558	vol/vol ubRMSD and a correlation coefficient of 0.55 against the in-situ soil moisture

559	measurements. ,It convergesing more apparently to the 1:1 line when compared with

560	validation result of the SPL2SMAP_S_V3 dataset in Fig. 6-(f). As with the area of

561	China, therefore, the currently developed product is generally superior to the global

562 SMAP/Sentinel combined SSM in terms of both coverage percentage and estimate
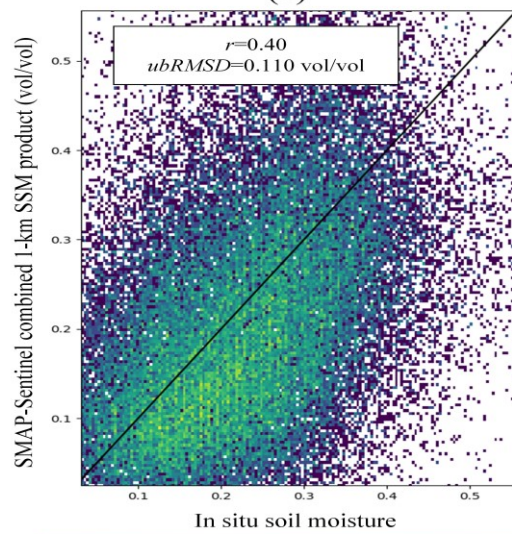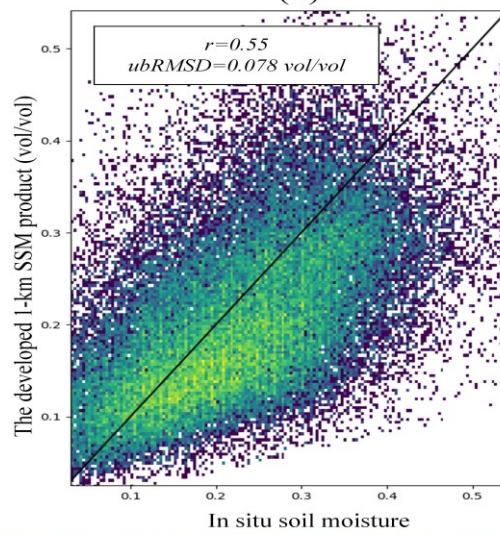
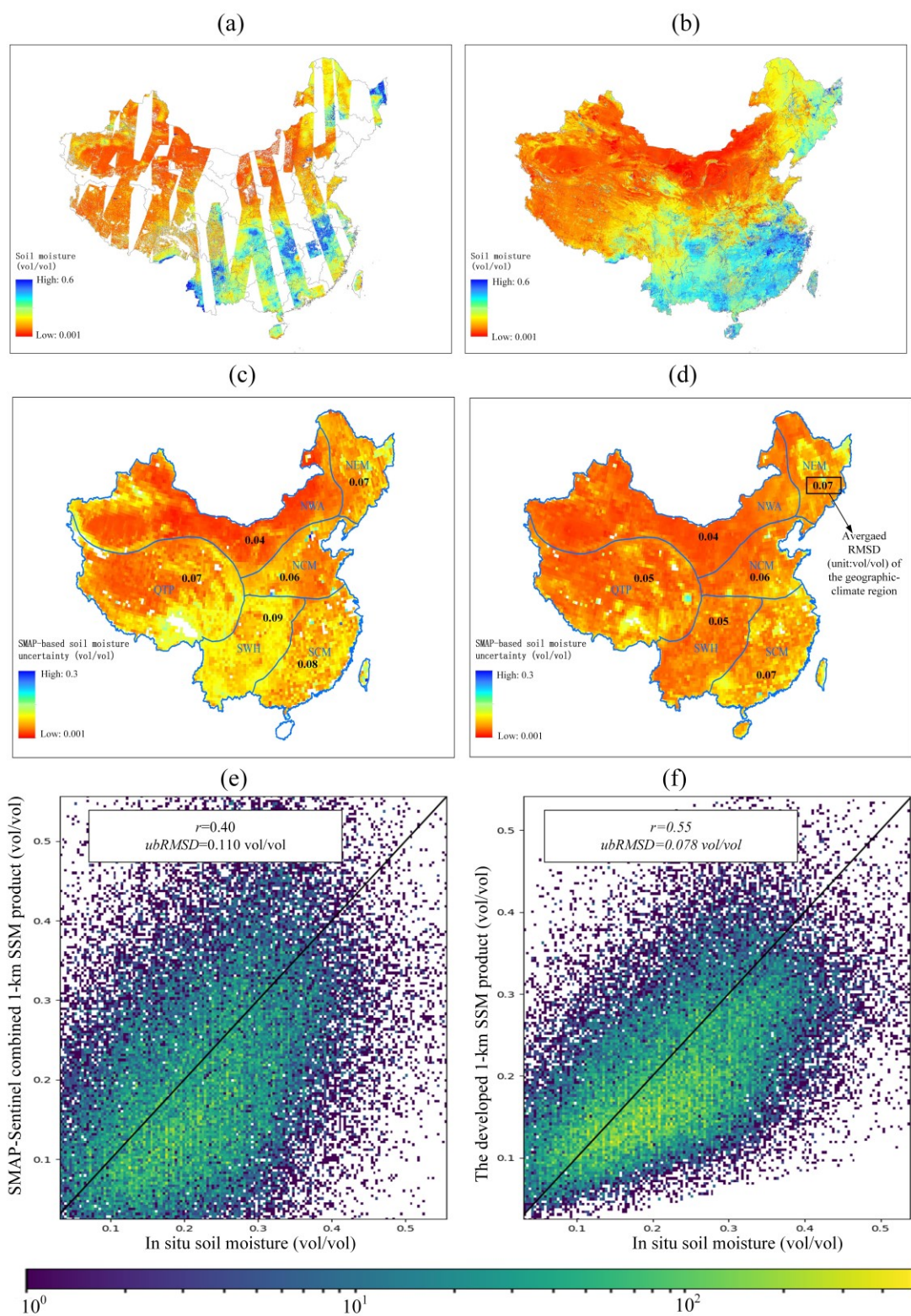**563** accuracy.

(a)
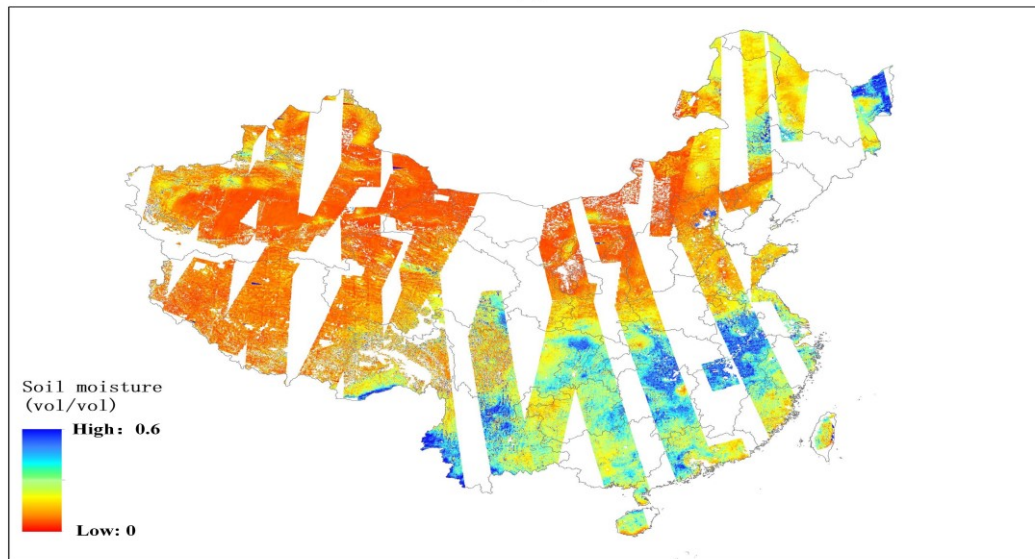
(b)
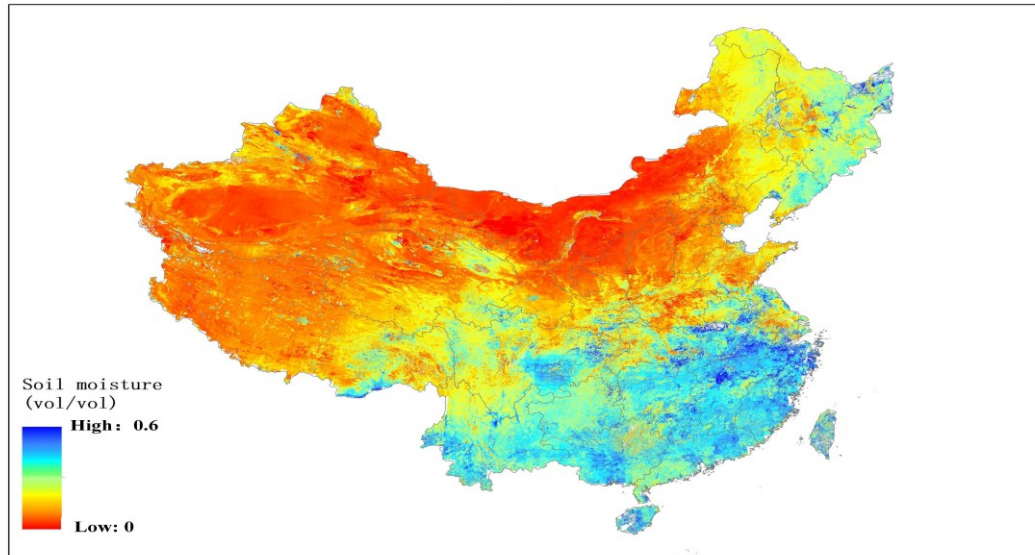
(c)

(d)

Fig. 6 Comparison results between the currently developed 1-km SSM product and the SMAP/Sentinel

combined 1-km SSM (SPL2SMAP_S_V3). (a) SPL2SMAP_S_V3 SSM images over China at about

6:00 a.m. systhesized by 6 continous dates from ~~April~~ May 26, 2018 to ~~May~~ April ~~3~~1, 2018. (b) The

SSM image at 1:30 a.m. of ~~May~~April 29, 2018 from the currently developed product. (c) Spatial

uncertainty (RMSD) map of the SPL2SMAP_S_V3 product against SMAP radiometer-based SSM

retrievals at the 36-km pixel scale over China for years of 2017, 2018, and 2019. (d) Same to (c) but for

validaiton of the currently developed SSM product. The black numbers in each of the geographic-

climate regions indicate averaged uncertainty (RMSD, unit: vol/vol) of the region. (e~~c~~) Validation

results of the SPL2SMAP_S_V3 product against in-situ soil moisture measurements over China for

years of 2017, 2018, and 2019. The black solid line is the 1:1 line. (f~~d~~) Same to (e~~c~~) but for validaiton

of the currently developed SSM product.

In Fig. 7~~Fig. 6~~, we display the cumulative distribution frequency of coverage

percentages of the downscaled SSM product and of the original PM NN-SM product

for each season. We should be noted that in this statistical scheme, pixels identified as

static water body by the MODIS MCD12Q1 land cover type product were not

considered in the denominator of the coverage percentage. Besides, the gap time

between the respective on-orbit period of AMSR-E and of AMSR-2 (from October

2011 to June 2012, during which there are no effective observations from the PM NN-

SM product) were also excluded. It is apparent that in Fig. 7~~Fig. 6~~-(b) and -(c), almost

all downscaled daily SSM images over the 16-17 years have achieved a coverage

percentage higher than 85%. In comparison, the majority of the PM NN-SM daily

images have their coverage percentages below 80% over the study region, primarily

due to the PM-seam gaps particularly existing in low latitudes (see Section 2.2.2). In

Fig. 7~~Fig. 6~~-(a) and -(d), the percentages of effective pixels in both the PM and the

downscaled SSM images are far lower than their counterparts in the other two

591 subfigures. This is mainly ascribed to extreme meteorological conditions including

592 snow, ice, and frozen soils that are typically persistent throughout most of these

593 specified months in the northwestern regions of China. Such conditions can impede

594 reliable estimates of SSM based on all satellite remote sensing techniques in the current

595 time. The above inter-seasonal differences on data coverage are also reflected in Fig.

596 8Fig. 7 in another manner based on presenting the spatial distributions of number

597 percentages of available dates in each three-month period.

598
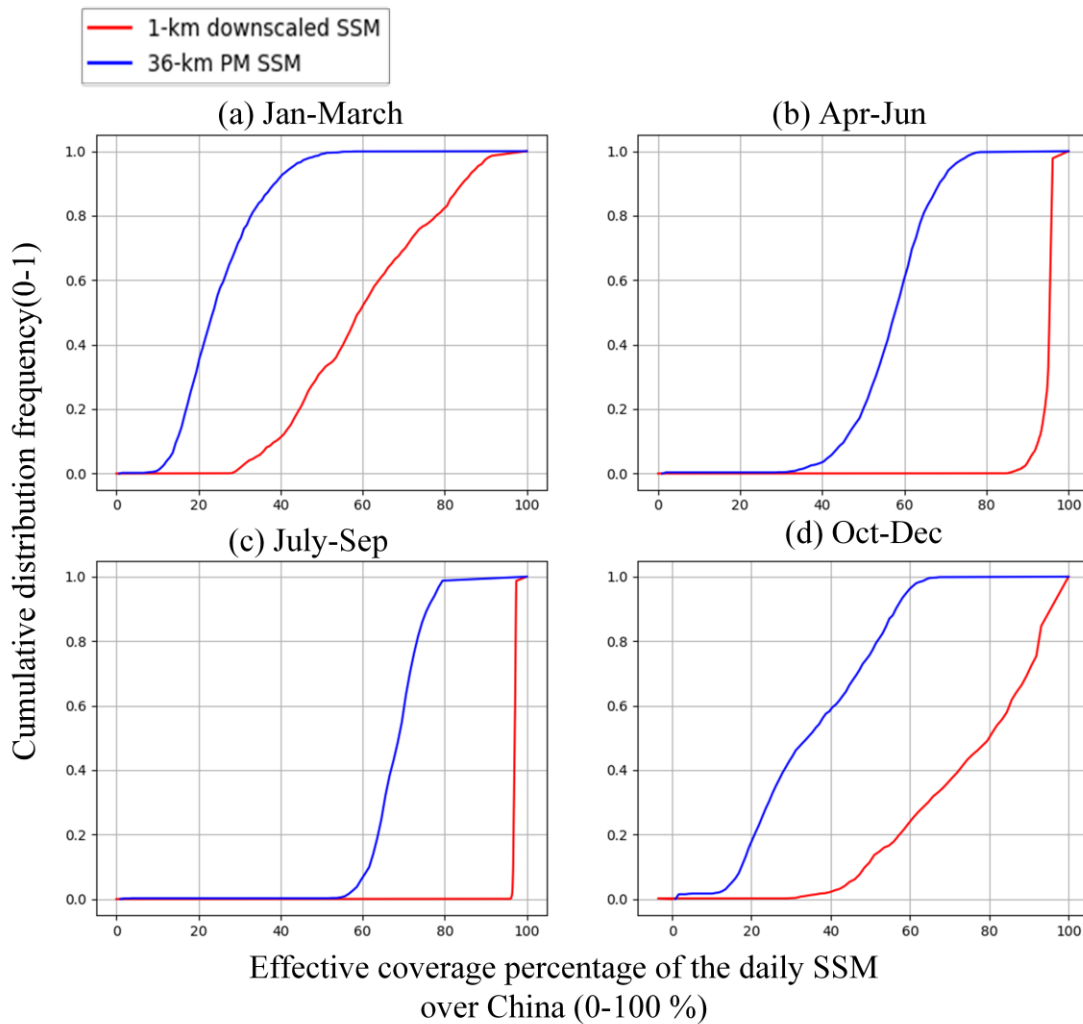
599   Fig. 7 Cumulative distribution frequency of our proposed SSM product against the original 36-km SSM

600   product for different seasons. The period between October 2011 and June 2012 is excluded in the

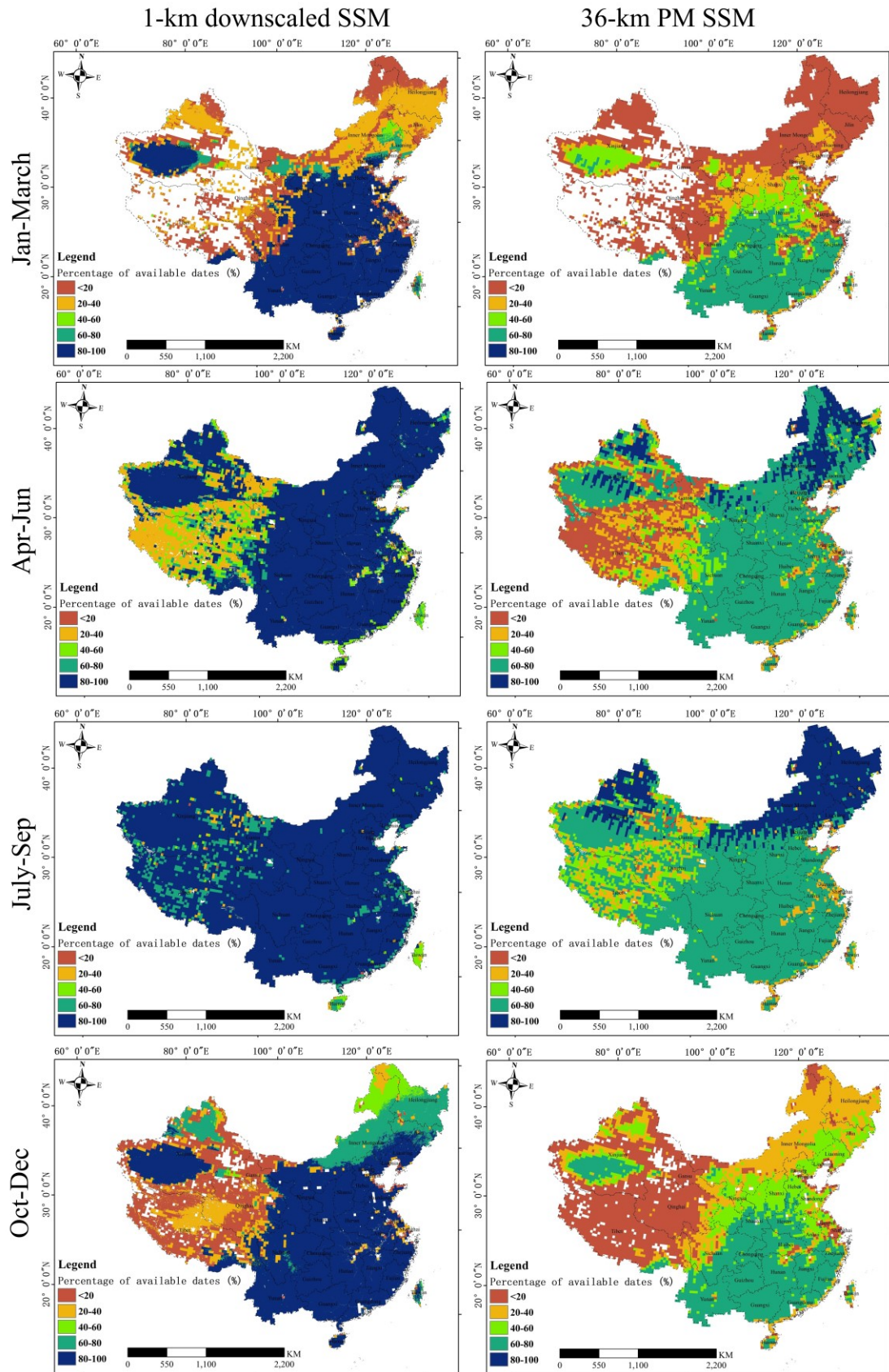**601**   current statistics.

602

603

Fig. 8 Spatial distributions on percentage of day numbers with available estimates for the currently

developed 1-km SSM product and the original 36-km PM data during 2003-2019. The four different

606   periods (i.e., January-March, April-June, July-September, October-December) of a year are treated

607   respectively. The period between October 2011 and June 2012 is excluded.

608   The techniques behind coverage improvement of the downscaled SSM (against

609   PM and optical data inputs) can be categorized into two classes, i.e. cloud gap-filling

610   of the input optical datasets (see Section 2.2.1), as well as the filling of downscaled

611   SSM in PM-seam gaps (see Section 2.2.2). Table 2Table 2 reports the specific

612   validation results (using averages of ground measurements at all stations) of

613   downscaled SSM in these coverage-improved conditions, relative to that generated

614   without using any coverage improvement technique, in order to evaluate the propagated

615   effect of such techniques on the final product. The very limited difference for ubRMSD

616   values (0.053 vol/vol versus 0.056 vol/vol) between cloudy and clear-sky conditions

617   suggest that the 1-km SSM estimates from our final product are generally compatible

618   between cloudy and clear-sky conditions. The downscaled SSM estimated for regions

619   of PM-seam gaps have a slightly worse (but still acceptable) accuracy, considering its

620   ubRMSD of 0.059 vol/vol compared to the 0.052 vol/vol ubRMSD of the PM-observed

621   1-km pixels. In summary of Fig. 7Fig. 6 and Table 2Table 2, the currently developed

622   product has achieved a substantially improved spatial coverage against the original

623   remote sensing input datasets, whilst successfully preserved the SSM downscaling

624   accuracy of the observation-covered pixels at the same time.

Table 2 Comparisons between validation results for pixels under coverage-improved regions and

for pixels under remote-sensing-observation-covered regions.

Evaluation metric[*]	Comparison between cloudy and clear-sky conditions		Comparison between passive microwave (PM) observed regions and regions of PM-seam gaps	
	Clear-sky condition	Cloudy condition	PM-observed regions	PM-seam gaps
ubRMSD (vol/vol)	0.053	0.056	0.052	0.059
Correlation coefficient	0.49	0.47	0.49	0.44

[*]All evaluation metrics in this column indicate the average of all available stations

3.4. Discussion

3.14.1Uncertainty on SSM evaluation between satellite- and ground- scales

In this study, we made evaluations on remote sensing SSM products at different spatial resolutions, using measurements from 2000+ stations provided by the national-level soil moisture observation network of China as standard benchmark. Through the evaluations, a ubRMSD of 0.074 vol/vol is reported for the original 36-km NN-SM SSM product (Fig.A1-b). We notice that this result is considerably poorer if compared with another previous evaluation campaign targeting at the same product (Yao et al., 2021), which achieved a global RMSE (RMSD) of 0.029 vol/vol. However, this difference is not unexpected because the two campaigns were carried out in different regions of the world. Also, that particular study (Yao et al., 2021) was conducted based on completely different ground soil moisture observations provided by the International

641 Soil Moisture Network (ISMN) (Dorigo et al., 2021). Compared to the observation

642 network employed in this study, the observation sites of ISMN are more intensively

643 distributed as an "integrated soil moisture station" so as to provide spatially average

644 soil moisture within a grid of tens of kilometers. In this regard, we admit that the ISMN

645 is generally more professional in evaluating satellite PM-based SSM retrievals at a

646 coarser resolution. But on the other hand, only a few (≤4) of such "integrated stations"

647 have been set up sporadically within China, making the ISMN data much less

648 representative of our study region compared with the national-level soil moisture

649 network of China exploited by our current study.

650 Although the higher RMSD of the national-level soil moisture network of China

651 may indicate larger measurement uncertainty than the ISMN, the negative influence

652 that might be imposed on our study purpose should be inconsequential. This is because

653 we focus more on the relative validation performance of different SSM products, rather

654 than on the absolute value of any evaluation metric including ubRMSD and correlation

655 coefficient calculated against ground measurements. Specifically, the 1-km downscaled

656 SSM obtained an average ubRMSD of about 0.054 vol/vol among different stations

657 according to Fig. 4Fig. 4-(b). Besides, result of the evaluation in Fig. 6Fig. 5-(d) based

658 on combination of multi-station ground measurements shows a global ubRMSD of

659 0.078 vol/vol for this product. Overall, the above-mentioned results can be identified

660 as at least comparable to the global (multi-station based) ubRMSD of 0.074 vol/vol of

661 the original NN-SM data as they are evaluated against the same benchmark. Therefore,

662 conclusion is safely drawn that the currently developed product preserves the retrieval

663   accuracy of the coarse-resolution NN-SM data, whilst improving the spatial

664   representativeness of the latter product substantially according to the mostly positive

665   G_{down} values in Fig. 4Fig. 4-(a).

666   Moreover, one may also argue that the r-value of 0.55 for the currently developed

667   product in Fig. 6Fig. 5-(d) is not sufficiently high compared with several previous

668   studies (Wei et al., 2019; Sabaghy et al., 2020) obtaining r-values above 0.7 for

669   temporal analysis of satellite remote sensing soil moisture. However, we should be

670   noticed that these previous studies have conducted analyses respectively at the temporal

671   and the spatial dimensions. Based on their results, the spatial analysis typically derived

672   lower r-values ($<$0.4) compared to that at the temporal dimension. This is probably

673   because the heterogeneity degree of remote sensing pixels can vary significantly across

674   different sites. Since the evaluation in Fig. 5-(d) was deployed at the 'spatio-temporal'

675   integrated dimensions, such an r-value is expected. This is also close to the global r-

676   value of 0.6 for validation of the coarse-resolution NN-SM product as reported in Yao

677   et al. (2021).

678   ## 3.24.2 Uncertainty on cloud gap-filling and validations of LST

679   As mentioned in Section 2.2.1, LST gap-filled based on the STDF method was

680   used alone as one of the main input datasets for SSM downscaling under cloudy weather.

681   Although such LST inputs contain clear-sky bias from the real cloudy condition, it

682   performs better in driving the SSM downscaling model compared with its bias-adjusted

683   counterpart (see Section Appendix-B for details). The reason may be linked to one of

684   the basic theories behind our SSM downscaling methodology, i.e. the "universal

685　triangle feature space (UTFS)" theory (Carlson et al., 1994). In the UTFS, clear-sky

686　LST is employed to implicitly quantify the surface soil wetness degree as it correlates

687　with the dynamics of soil evaporative efficiency and soil thermal inertia when

688　vegetation cover density is fixed. Under cloudy conditions, however, the satellite

689　observed LST is subjected to not only surface soil property, but also to that related to

690　cloud insulation effect from solar incoming radiation and ground long wave outgoing

691　radiation. As a result, the actual relationship between SSM and cloudy LST could be

692　much more complicated than the one that has been described by the UTFS-based SSM

693　downscaling model (i.e. Equation-2). In comparison, LST generated by the STDF alone

694　for assumed clear-sky conditions, as is free from interference of cloud, would be a

695　comparatively more competent input variable for driving the UTFS-based SSM

696　downscaling model under non-rainy clouds. This is especially the case for thin and

697　short-time clouds with marginal direct feedbacks on surface soil wetness.

698　　　However, we admit that the STDF-filled LST under rainy clouds is also not suitable

699　for our study purpose. This may explain the slightly higher RMSD for SSM under cloud

700　based on STDF-filled LST (0.056 vol/vol) compared to that under real clear sky (0.053

701　vol/vol), as shown in Table 2Table 2. In reality, the actual negative influence of cloud

702　on the final SSM product may be even more serious than indication from the above

703　RMSD difference (i.e. 0.056-0.053 = 0.003 vol/vol), due to the portion of

704　"clear/cloudy-weather-mixed" spatial windows during the fitting process of the

705　downscaling model. In these windows, uncertainty in cloud gap-filled LST may affect

706　accuracy of the fitted model coefficients and thus deteriorate the final SSM estimates

43

707 in clear-sky pixels within the same window. Consequently, the above RMSD difference

708 has been more or less underestimated. Despite all of above, in our study area of China

709 we regard the STDF-filled LST as a more optimal proxy of heat flux for estimating

710 SSM under clouds, compared to the bias-adjusted LST. On the other hand, future efforts

711 are encouraged to further clarify the mechanical relationships between STDF-

712 filled/bias-adjusted LST and soil wetness degree under clouds.

713 　　Different from a number of previous studies (Jiménez et al., 2017; Dowling et al.,

714 2021; Yang et al., 2019) validating satellite thermal-infrared-based LST based on

715 longwave radiation observations made at footprint-level observation stations (e.g. flux

716 towers), our study has used 0-cm top ground temperatures as the primary benchmark

717 for this validation campaign instead. Similar to that for SSM validation, the most crucial

718 motivation driving such an experimental design is the significantly intensive

719 distribution of the meteorological stations compared to the very limited number of

720 active and effective flux towers available in China. It is noted that these measurement

721 devices at all of the meteorological stations are required to have been instrumented

722 under open environmental conditions with relatively lower fraction of tall trees and

723 water bodies, in order to conduct efficient monitoring at the physics of near-surface air.

724 This can also be reflected in Fig.4-(c), which reveals no stations built within forest

725 covers. Moreover, as we only focus on the mid-night scenario when the states of all

726 land observations are "most stable" during one diurnal cycle, uncertainties due to the

727 possible temperature inconsistency between bare ground surface and high tree surface

728 as well as due to the temporal mismatch (from about 1:30 to 2:00 A.M.) should have

729 margional effect on our results. We have carried an extra test that can confirm this

730 discussion, with the detailed procedures described in Section Appendix-C.

4.3 Major novelty, unique profit, and future prospect of the developed product

733 Compared with the widely known active/passive microwave combined SSM

734 product (e.g. the SPL2SMAP_S_V3) and other PM/optical-data combined counterparts

735 which were also published recently but at the monthly scale (Meng et al., 2021), the

736 major novelty of the currently developed product mainly lies in the fact that it has

737 achieved progress on all of the three crucial dimensions of satellite remote sensing,

738 including the temporal revisit cycle (daily), the spatial resolution (1-km), and the quasi-

739 complete coverage under all-weather conditions. To our knowledge, this has rarely been

740 achieved by previously developed satellite soil moisture product at regional scales. For

741 realization of the above-mentioned progresses, we have fused the SSM downscaling

742 framework with other techniques including cloud gap-filling of thermal infrared LST,

743 MWS-based temporal filtering of vegetation indices, as well as reconstruction of seams

744 between neighboring PM swaths in low latitudes. The final SSM estimates under cloudy

745 conditions and intersected with the PM-seam gaps were specially validated against the

746 rest estimates under clear sky and in the regions covered by PM observations,

747 respectively (Table 2Table 2). The comparable performances among all treatment

748 groups herein confirm that the accuracy of the product is stable and consistent among

749 all weather conditions.

With improvement achieved at the three dimensions, unique profit of the currently developed product can be taken by subsequent studies and various industrial applications. For example, the capability of this product can be investigated on capturing the short-term anomaly of local hydrological signals as well as improved monitoring on drought disasters, which used to be investigated mainly at a coarser resolution by PM SSM (Scaini et al., 2015; Champagne et al., 2011; Albergel et al., 2012). For another, taking advantage of its all-weather daily time series, the product can be utilized together with precipitation data to isolate and quantify the anthropic influence on regional water resources from the natural hydrological dynamics. Examples of such anthropic signals include agricultural irrigation activities, as well as finer-scale information on agricultural crops which was previously interpreted based on PM-driven techniques (Song et al., 2018). In addition, we should realize the important role of soil moisture as a constraint for accurate estimation of surface evapotranspiration and runoff (Zhang et al., 2020; Zhang et al., 2019). Therefore, the profit of this product can be further enhanced if coupled with land-atmosphere coupled models to produce new insights into water-cycle processes of earth surface at a finer spatio-temporal scale.

There are still some limitations on our current ~~version of this developed~~ product to be further improved ~~at the following aspects i~~In the future. First, there may exist~~is concerned with~~ the 'mosaic effect' at the original PM (36-km) pixel edge. As mentioned in Section 2.2.2, we have used a parameter of 'spatial square window (*ws*)' in Equation-(3) to ~~decline~~minimize this negative effect ~~to the best of our ability~~. H~~,~~owever, it is still

difficult to utterly avoid such negative ~~influence~~effect. This is a challenge for ~~in~~ all existing SSM downscaling ~~studies~~methods (Molero et al., 2016; Stefan et al., 2020; Peng et al., 2016), especially considering the large spatial scale of our study and all uncertainties discussed in Sections 4.1 and 4.2. Besides, other negative influences can be imposed by potential imperfections identified from the original PM product, e.g. from PM SSM retrievals in the QTP region with complicated topography, melt snow or partially frozen soils that cannot been completely screened out by the original product flag in winter. For these extreme conditions, the accuracy of the downscaled SSM may need further validation campaigns like field surveys and experiments, based on which the data quality flag can be better built for the product's futural version.

~~T~~the methodological framework proposed in this paper is prospective to be universally applied in other regions of the world to serve for better monitoring of the global surface wetness in the following studies. If applied in continental and global scales, however, the current process for gap-filling of PM seams may require further attention and improvement. In this study, SSM in regions intersected with PM-seam gaps were estimated using TB observations from PM swaths at neighboring dates (see Equation-5). Although the errors in the PM-seam gaps over China as reported by Table 2~~Table 2~~ are only slightly larger compared to the PM-covered regions, they cannot be ignorable completely and may leave extra concern on the universality of this technique, especially in the low latitudinal tropical regions where the effect of PM-seam gap is more apparent than in our study area. Besides, another imperfection of this data product lies in the gap period between AMSR-E and AMSR-2. Considering the different

794 systematic error patterns of various PM SSM products, we did not generate downscaled

795 SSM based on other PM products (e.g. the SMOS SSM product) during this period but

796 just left the period as null values. We suggest a more rigorous and universal inter-

797 calibration framework on different PM SSM products to be developed in the future for

798 a long-term consistent 1-km downscaled SSM dataset.

4.5.   Conclusions

800   This paper describes the main technical procedures of a recently developed remote

801 sensing surface soil moisture (SSM) product over China covering the recent ten years

802 and more. Based on combination of passive microwave SSM downscaling theory and

803 other related remote sensing techniques, the product achieves multi-dimensional

804 distinctive features including 1-km resolution, daily revisit cycle, and quasi-complete

805 all-weather coverage. These were rarely satisfied completely by other existing remote

806 sensing SSM product at regional scales. Validations were conducted against

807 measurements from 2000+ automatic soil moisture observation stations over China.

808 Overall, an average ubRMSD of 0.054 vol/vol across different stations is reported for

809 the currently developed product. The mostly positive G_{down} values show this product

810 has significantly improved spatial representativeness against the 36-km PM SSM data

811 (a major source for downscaling). Meanwhile, it generally preserves the retrieval

812 accuracy of the 36-km data product. Moreover, additional validation results show that

813 the currently developed product surpasses the widely used SMAP-sentinel combined

814 global 1-km SSM product, with a correlation coefficient of 0.55 achieved against that

815 of 0.40 for the latter product. At the regional scale, time series patterns of our developed

816 data product are highly correlated with that of the widely recognized SMAP radiometer-

817 based SSM for all geographic settings. The methodological framework for product

818 generation is promising to be applied at the continental and global scales in the future,

819 and the product is potential to benefit various research/industrial fields related to

820 hydrological processes and water resource management.

821

Appendix

A. Evaluation on different PM SSM products

We have made evaluations on the various AMSR-based SSM products (as shown in Table 1~~Table 1~~) covering the recent 10 years or longer, based on our soil moisture observation network all over China. The ~~L band based~~ SMAP radiometer-based SSM dataset, as described in Section 2.1.4, was also evaluated as a reference. The evaluation period covers the three years of 2017, 2018, and 2019. All AMSR-based 25-km grids were re-set to the SMAP 36-km grid system using the nearest resampling method. Only grids that contain equal or more than 4 soil moisture measurement stations were employed, in which, the grid-based PM SSM estimate was compared with average of measurements from all interior stations. Finally, 53 grids were selected, as shown by the green color in Fig.A1-(g). For AMSR-based products, only the mid-night descending datasets were evaluated, whist for the SMAP product, our evaluation only focused on its descending mode in the early morning.

As manifested by Fig.A1-(a) to -(f), the selected SSM product in the current study, i.e., the NN-SM product has an unbiased RMSD of 0.074 vol/vol and a correlation coefficient of 0.49. This obviously outperforms the other three traditional AMSR-based SSM products (i.e. JAXA-AMSR, LPRM-AMSR, and UMT-AMSR products) and is only inferior to the SMAP SSM retrievals, whilst the later only covers the latest period since 2015. As far as CCI data are concerned, it has a similar performance against the selected NN-SM in general. Nevertheless, the region marked by red circle in Fig.A1-

843 (c) indicates that CCI estimates have a considerably larger proportion of overestimated

844 anomalies. But overall, the primary reason that we have abandoned CCI but selected

845 NN-SM is because the latter can provide a higher coverage fraction of valid pixels in

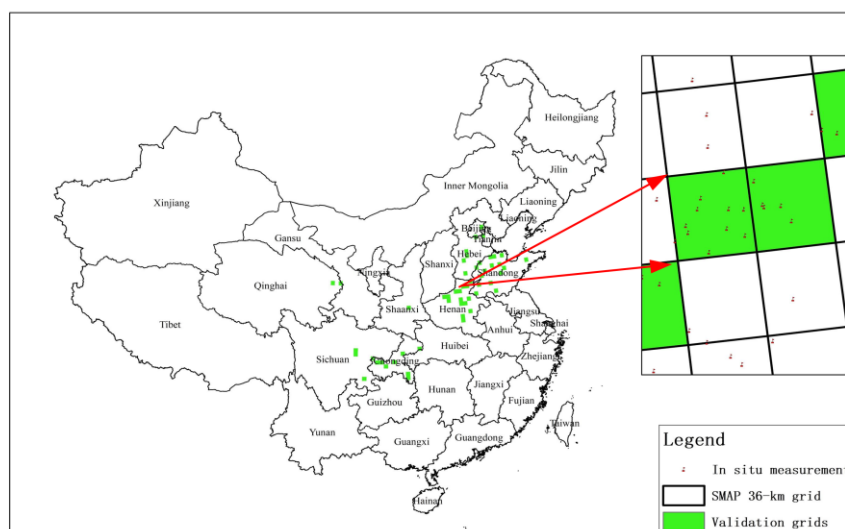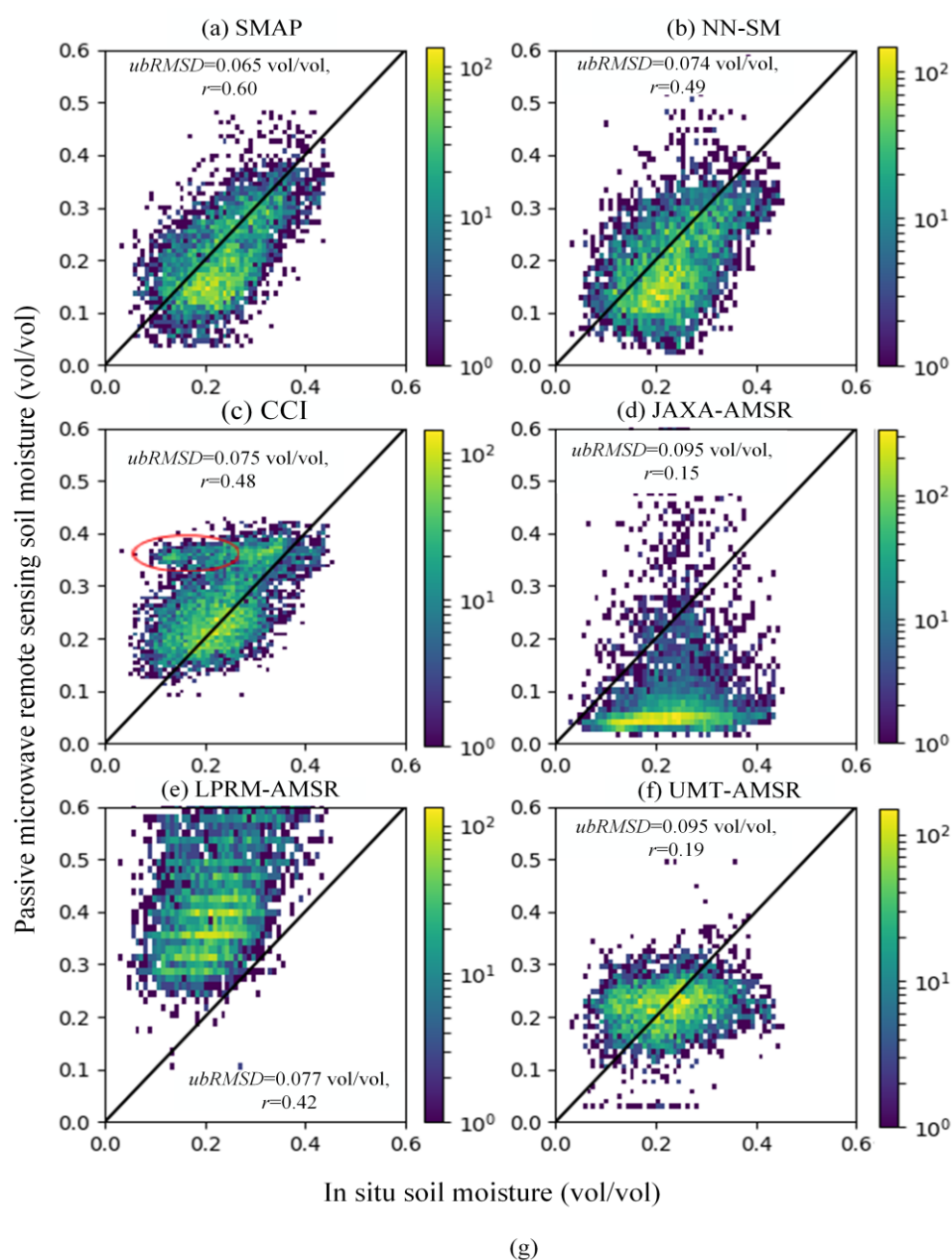846 our study region, as has been stated in Section 2.1.1.

(g)

Fig. A1 (a)-(f) Comparison of different PM SSM products (as reported in Table 1~~Table 1~~) against the

in situ SSM measurements in China. (g) Locations of the 36-km EASE-GRID-projection based pixels

used for this comparison campaign.

B. Evaluation on the influence of bias adjustment for reconstructed 'clear-sky' LST under cloud

In Section 2.2.2, we have emphasized that the gap-filled LST for cloudy pixels

reflects the theoretical surface temperature of that pixel under a hypothetical clear-sky

condition. As this cloud gap-filled LST would suffer from a possible bias against the

real surface temperature under cloud (Dowling et al., 2021), we made an additional

experiment regarding to further improvement of this cloud gap-filled LST. The follow-

up step for bias adjustment of this hypothetical clear-sky LST (but actually under

cloudy conditions), as expounded in Section 4.2 of Dowling et al. (2021), was

conducted herein using remote sensing and in situ LST data over China but only in

2018. We illustrate the validation results for bias adjusted and non-bias adjusted LST

under cloudy conditions in Fig. A2-(b) and -(c), respectively. Similar to Fig. 3~~Fig. 3~~,

validation results for clear-sky LST of that year are also displayed (Fig. A2-(a)) for

comparison. The results generally show that the follow-up step is effective in reducing

the bias of the originally gap-filled 'clear-sky LST' under cloudy conditions (from -1.7

K to 0.4 K).

In the subsequent step, we substituted the original non-bias adjusted LST under

cloudy conditions with its bias adjusted counterpart, and used the latter as the input for

SSM downscaling. The general validation results of the downscaled SSM are illustrated

870   in Fig. A3 (similar to that presented in Fig. 4~~Fig. 4~~-a and -b). Contrary to the above-

871   analyzed Fig. A2, the bias adjusted cloudy LST with better gap-filling accuracies,

872   however, obtained inferior performance in SSM downscaling. This final validation

873   result, to some degree, confirms our assumption in Section 2.2.2 that the reconstructed

874   cloudy LST but for the hypothesized clear-sky condition is the better proxy of surface

875   moisture dynamics. But overall, as all LST estimates discussed herein are for the

876   midnight scenario (when the energy interaction between atmosphere and land surface

877   is relatively weak), the RMSD difference for different weather conditions in Fig.A2 is

878   expectedly marginal. As a consequence, the difference in ubRMSD of SSM in Fig.A3

879   can hardly be identified as 'very significant'. Therefore, we encourage further tests on

880   this conclusion in specific future studies to confirm its universality, especially for

881   situation of the 'morning to noon' time window.
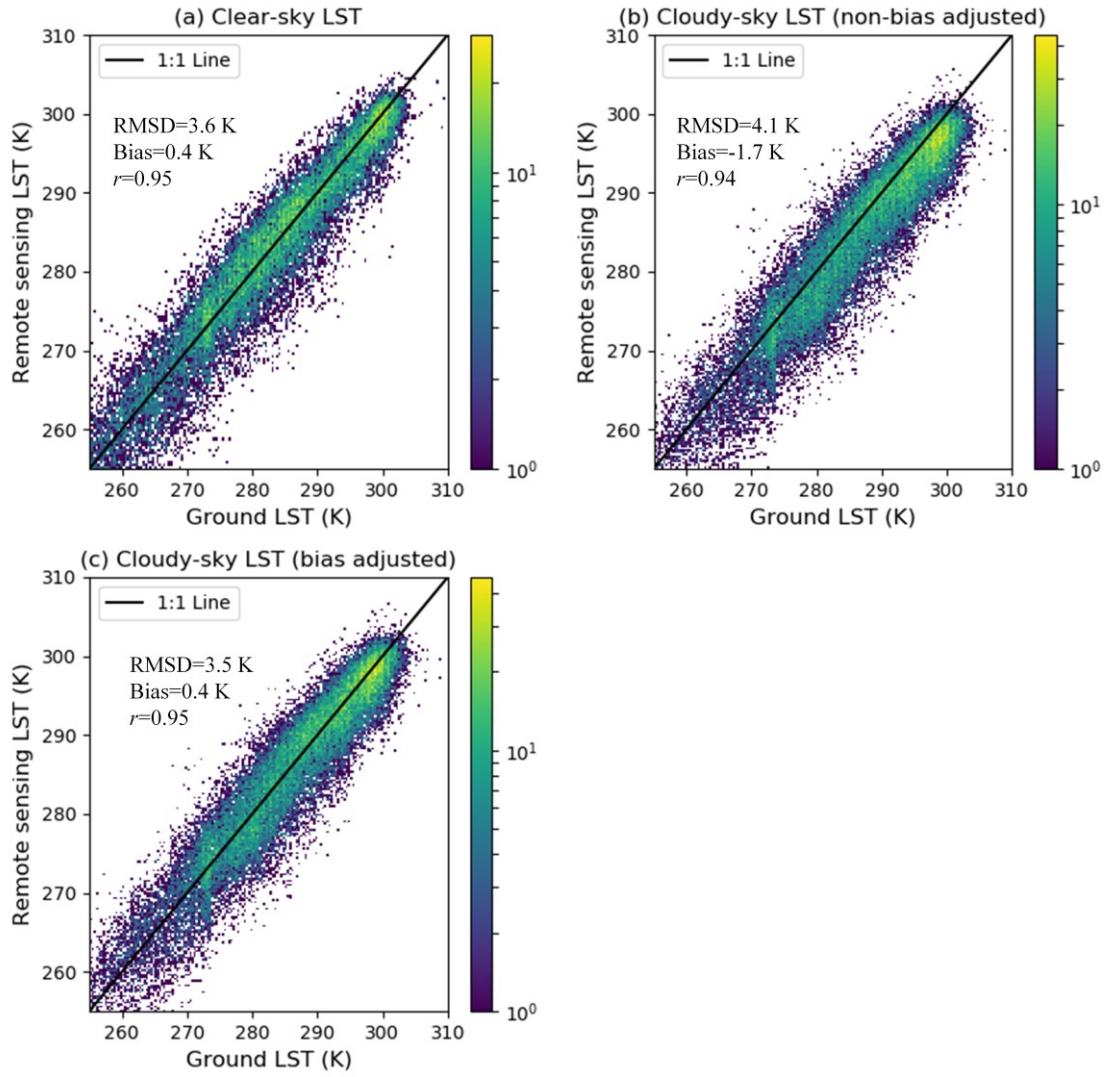
882

883

884   Fig. A2 Validation of the clear sky LST (a), reconstructed LST under cloud but with no passive-

885   microwave based bias adjustment (b), as well as the reconstructed LST under cloud with passive-

886   microwave based bias adjustment (c) respectively, based on the 0-cm ground temperature

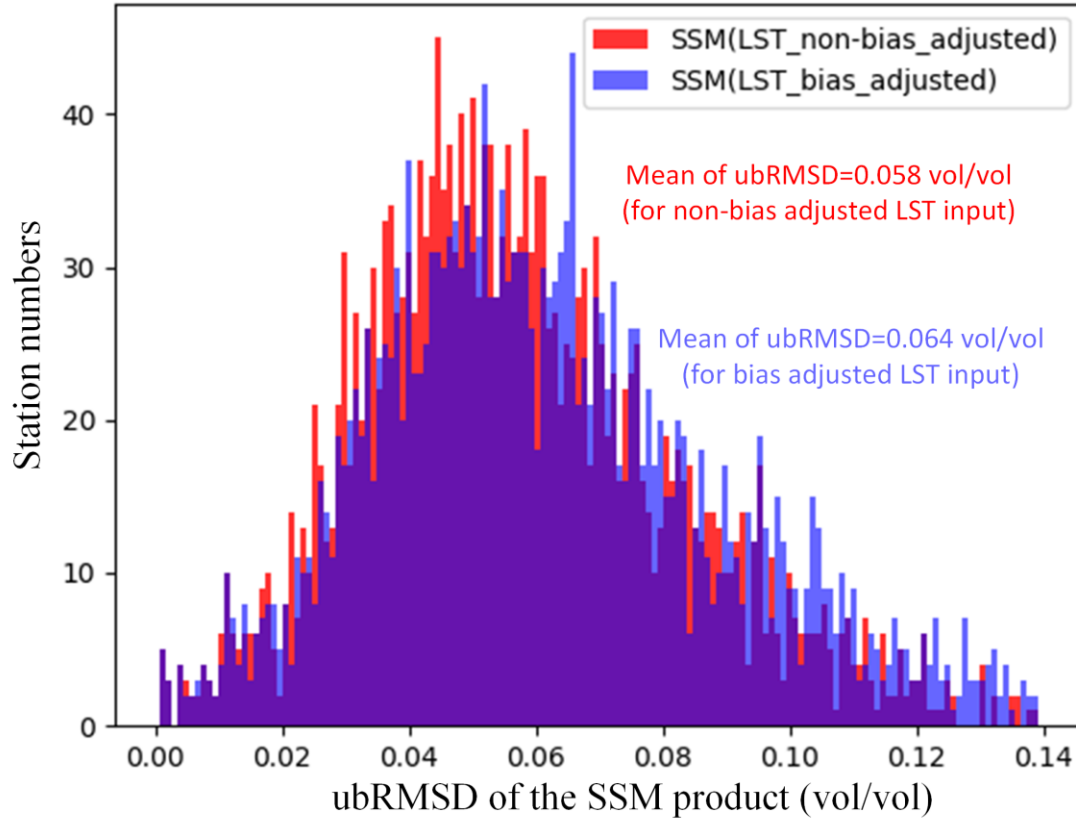**887**   measurements at meteorological stations.

888

Fig. A3 The statistical distribution of ubRMSD at different stations for SSM estimates driven by two

respective kinds of cloudy LST inputs.

C. Uncertainty test between 0-cm ground temperature observations and flux-tower-derived thermal infrared LST

We herein utilized 4 flux towers to calculate their footprint-level (about 500-1000 m) thermal infrared LST based on long wave radiation measurements, plus broad band emissivity data derived from the MODIS MYD21A1 product (MYD21A1N.V061). The 4 towers are all characterized by moderate or low vegetation (grassland) and are dispersedly located at different eco-regions of China, namely the towers of Changling, Huailai, Yakou, and Naqu (see the inset map in Fig.A4-b). Data from Changling are derived from the FLUXNET community (FLUXNET2015 Dataset - FLUXNET) in 2010.

901 Data from the other three towers are derived from the National Tibetan Plateau Data

902 Center, with data DOIs of http://dx.doi.org/10.11888/Meteoro.tpdc.271094 for Huailai

903 in 2018, http://dx.doi.org/10.11888/Meteoro.tpdc.270781 for Yakou in 2018, and

904 http://dx.doi.org/10.11888/Meteoro.tpdc.270910 for Naqu in 2016. These data have

905 been preprocessed by their providers to record the dynamics of those variables at a half-

906 hour interval. The algorithm for calculating LST based on flux-tower-derived long

907 wave radiation is inherited from Wang and Liang (2009). We first compared the flux-

908 tower-derived night-time LST estimates between 1:00-1:30 A.M. and 1:30-2:00 A.M..

909 As shown by Fig.A4-(a), the very slight RMSD of 0.72 K suggests that LST is generally

910 stable between 1:00 and 2:00 A.M. at night. In Fig.A4-(b), we also found marginal bias

911 and RMSD within 1 K between average flux-tower-derived LST of 1:00- 2:00 A.M.

912 and the corresponding 0-cm ground temperature at close meteorological sites (within 1

913 km and at 2:00 A.M.).

914   In Fig.A4-(c) we demonstrate time series for monthly average NDVI (derived as

915 in Section 2.2.1) at the 1-km pixels containing each of the four sites from 2003-2019.

916 Clearly, there are very rare cases with NDVI values exceeding 0.5, corroborating the

917 "open environmental conditions" met by the meteorological stations. In view of above,

918 it is feasible for our study to have used the 0-cm ground temperature at pixels of such

919 moderate to low vegetation covers as the evaluation benchmark of the satellite-derived
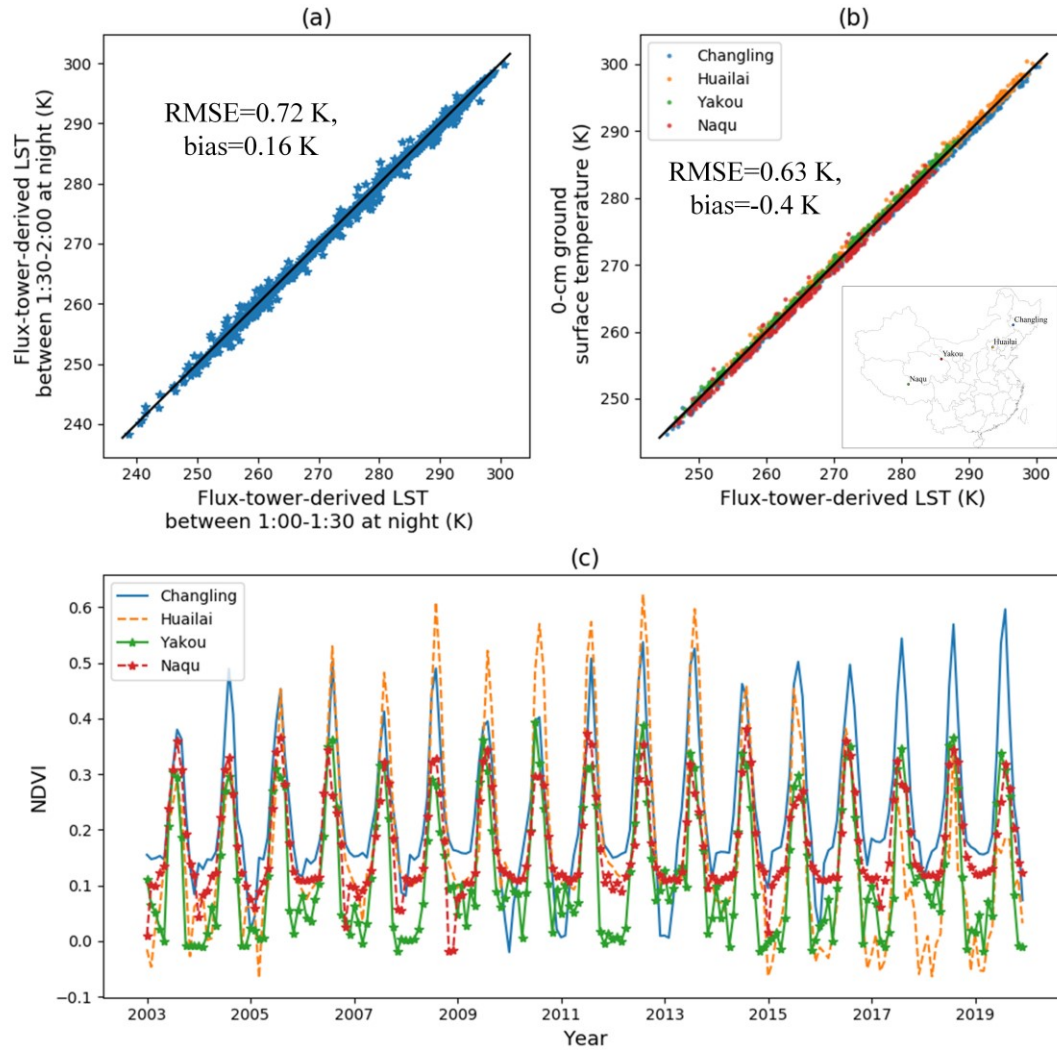
920 thermal infrared LST.

Fig. A4 (a) Comparison of LST between 1:00-1:30 A.M. and 1:30-2:00 A.M. for the four selected flux towers. (b) Comparison of flux-tower-derived LST averaged for 1:00-2:00 A.M. at the four towers and corresponding night-time 0-cm ground temperature at proximal meteorological stations. The inset map shows the location of the four flux towers. (3) Monthly NDVI time series for 1-km pixels containing each of the four flux towers.

Author contributions

Peilin Song and Yongqiang Zhang designed the research and developed the whole methodological framework; Peilin Song and Yongqiang Zhang supervised the processing line of the 1-km SSM product; Jianping Guo and Bingtong provided in situ

soil moisture data for validation; Peilin Song wrote the original draft of the manuscript;

Yongqiang Zhang, Jiancheng Shi, and Tianjie Zhao revised the manuscript.

Competing interests

Data availability

The published SSM dataset is available under the Creative Commons Attribution

4.0 International License at the following link:

http://dx.doi.org/10.11888/Hydro.tpdc.271762 (Song and Zhang, 2021b). This dataset

covers all of China's terrestrial area at a daily revisit frequency (about 1:30 A.M. at

local time) and a 1km spatial resolution from January 2003 to October 2011 and from

July 2012 to December 2019.

Acknowledgement

Financial support

References

Albergel, C., de Rosnay, P., Gruhier, C., Munoz-Sabater, J., Hasenauer, S., Isaksen, L., . . . Wagner, W.: Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sens. Environ., 118, 215-226, 10.1016/j.rse.2011.11.017, 2012.

Brunsdon, C., Fotheringham, A. S., and Charlton, M. E.: Geographically weighted regression: A method for exploring spatial nonstationarity, Geogr. Anal., 28, 281-298, 1996.

Busch, F. A., Niemann, J. D., and Coleman, M.: Evaluation of an empirical orthogonal function-based method to downscale soil moisture patterns based on topographical attributes, Hydrological Processes, 26, 2696-2709, 2012.

Carlson, T. N., Gillies, R. R., and Perry, E. M.: A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover, Remote sensing reviews, 9, 161-173, 1994.

Champagne, C., McNairn, H., and Berg, A. A.: Monitoring agricultural soil moisture extremes in Canada using passive microwave remote sensing, Remote Sens. Environ., 115, 2434-2444, 2011.

Chauhan, N. S., Miller, S., and Ardanuy, P.: Spaceborne soil moisture estimation at high resolution: a microwave-optical/IR synergistic approach, Int. J. Remote Sens., 24, 4599-4622, http://doi.org/10.1080/0143116031000156837, 2003.

Chen, Y., Yuan, H., Yang, Y., and Sun, R.: Sub-daily soil moisture estimate using dynamic Bayesian model averaging, J. Hydrol., 590, 125445, https://doi.org/10.1016/j.jhydrol.2020.125445, 2020.

Choi, M. and Hur, Y.: A microwave-optical/infrared disaggregation for improving spatial representation of soil moisture using AMSR-E and MODIS products, Remote Sens. Environ., 124, 259-269, http://doi.org/10.1016/j.rse.2012.05.009, 2012.

Das, N., Entekhabi, D., Dunbar, R. S., Kim, S., Yueh, S., Colliander, A., . . . Cosh, M.: SMAP/Sentinel-1 L2 Radiometer/Radar 30-Second Scene 3 km EASE-Grid Soil Moisture, Version 3 [dataset], https://doi.org/10.5067/ASB0EQO2LYJV, 2020.

Das, N. N., Entekhabi, D., Dunbar, R. S., Chaubell, M. J., Colliander, A., Yueh, S., . . . Thibeault, M.: The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product, Remote Sens. Environ., 233, 111380, https://doi.org/10.1016/j.rse.2019.111380, 2019.

979   den Besten, N., Steele-Dunne, S., de Jeu, R., and van der Zaag, P.: Towards Monitoring Waterlogging with Remote

980   Sensing for Sustainable Irrigated Agriculture, Remote Sens., 13, 2021.

981   Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., . . . Sabia, R.: The International

982   Soil Moisture Network: serving Earth system science for over a decade, Hydrol. Earth Syst. Sci., 25, 5749-5804,

983   10.5194/hess-25-5749-2021, 2021.

984   Dowling, T. P. F., Song, P., Jong, M. C. D., Merbold, L., Wooster, M. J., Huang, J., and Zhang, Y.: An Improved Cloud

985   Gap-Filling Method for Longwave Infrared Land Surface Temperatures through Introducing Passive Microwave

986   Techniques, Remote Sens., 13, 3522, 2021.

987   Du, J. Y., Kimball, J. S., and Jones, L. A.: Passive microwave remote sensing of soil moisture based on dynamic

988   vegetation scattering properties for AMSR-E, IEEE Trans. Geosci. Remote Sens, 54, 597-608, 2016.

989   Duan, S. and Li, Z.: Spatial Downscaling of MODIS Land Surface Temperatures Using Geographically Weighted

990   Regression: Case Study in Northern China, IEEE Trans. Geosci. Remote Sens, 54, 6458-6469,

991   http://doi.org/10.1109/TGRS.2016.2585198, 2016.

992   Entekhabi, D., Reichle, R. H., Koster, R. D., and Crow, W. T.: Performance Metrics for Soil Moisture Retrievals and

993   Application Requirements, J. Hydrometeorol., 11, 832-840, 10.1175/2010jhm1223.1, 2010a.

994   Entekhabi, D., Das, N., Kim, S., Jagdhuber, T., Piles, M., Yueh, S., . . . Martínez-Fernández, J.: High-Resolution

995   Enhanced Product based on SMAP Active-Passive Approach and Sentinel 1A Radar Data, AGU Fall Meeting Abstracts,

996   H24C-08,

997   Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., . . . Van Zyl, J.: The Soil

998   Moisture Active Passive (SMAP) Mission, Proc. IEEE, 98, 704-716, http://doi.org/10.1109/JPROC.2010.2043918, 2010b.

999   Fang, B. and Lakshmi, V.: Passive Microwave Soil Moisture Downscaling Using Vegetation and Surface Temperatures,

1000   Vadose Zone J, 12, 1712-1717, 2013.

1001   Fang, B., Lakshmi, V., Bindlish, R., and Jackson, T.: AMSR2 Soil Moisture Downscaling Using Temperature and

1002   Vegetation Data, Remote Sens., 10, 2018.

1003   Fang, B., Lakshmi, V., Bindlish, R., Jackson, T. J., Cosh, M., and Basara, J.: Passive Microwave Soil Moisture

1004   Downscaling Using Vegetation Index and Skin Surface Temperature, 2013.

1005   Fujii, H., Koike, T., and Imaoka, K.: Improvement of the AMSR-E Algorithm for Soil Moisture Estimation by

1006   Introducing a Fractional Vegetation Coverage Dataset Derived from MODIS Data, Journal of the Remote Sensing Society

1007   of Japan, 29, 282-292, 2009.

1008   Im, J., Park, S., Rhee, J., Baik, J., and Choi, M.: Downscaling of AMSR-E soil moisture with MODIS products using

1009   machine learning approaches, Environ Earth Sci, 75, 1-19, http://doi.org/10.1007/s12665-016-5917-6, 2016.

1010   Ines, A. V. M., Das, N. N., Hansen, J. W., and Njoku, E. G.: Assimilation of remotely sensed soil moisture and vegetation

1011   with a crop simulation model for maize yield prediction, Remote Sens. Environ., 138, 149-164, 10.1016/j.rse.2013.07.018,

1012   2013.

1013   Jeffrey, P., Walker, Paul, R., and Houser: A methodology for initializing soil moisture in a global climate model:

1014   Assimilation of near-surface soil moisture observations, Journal of Geophysical Research Atmospheres, 2001.

1015   Jiménez, C., Prigent, C., Ermida, S. L., and Moncet, J. L.: Inversion of AMSR‐E observations for land surface

1016   temperature estimation: 1. Methodology and evaluation with station temperature, Journal of Geophysical Research:

1017   Atmospheres, 2017.

1018   Jing, Z. and Zhang, X.: A soil moisture assimilation scheme using satellite-retrieved skin temperature in meso-scale

1019   weather forecast model, Atmos Res, 95, 333-352, 2010.

1020   Jones, L. A., Kimball, J. S., Podest, E., McDonald, K. C., Chan, S. K., and Njoku, E. G.: A method for deriving land

1021   surface moisture, vegetation optical depth, and open water fraction from AMSR-E, IEEE IGARSS 2009., Cape Town,

1022   South Africa, 2009, III-916-III-919, http://doi.org/10.1109/IGARSS.2009.5417921,

1023   Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., . . . Zhang, K.: Recent decline in

1024   the global land evapotranspiration trend due to limited moisture supply, Nature, 467, 951-954, 10.1038/nature09396, 2010.

1025   Kim, J. and Hogue, T. S.: Improving spatial soil moisture representation through integration of AMSR-E and MODIS

1026   products, IEEE Trans. Geosci. Remote Sens, 50, 446-460, http://doi.org/10.1109/TGRS.2011.2161318, 2012.

1027   Koike, T., Nakamura, Y., Kaihotsu, I., Davva, G., Matsuura, N., Tamagawa, K., and Fujii, H.: Development of an

1028   Advanced Microwave Scanning Radiometer (AMSR-E) algorithm of soil moisture and vegetation water content (written

1029   in Japanese), Annual Journal of Hydraulic Engineering, 48, 217-222 2004.

1030   Komatsu, T. S.: Toward a Robust Phenomenological Expression of Evaporation Efficiency for Unsaturated Soil

1031   Surfaces, Journal of Applied Meteorology, 42, 1330-1334, 10.1175/1520-0450(2003)042<1330:Tarpeo>2.0.Co;2, 2003.

1032   Kong, D., Zhang, Y., Gu, X., and Wang, D.: A robust method for reconstructing global MODIS EVI time series on the

1033   Google Earth Engine, Isprs J Photogramm, 155, 13-24, 2019.

1034   Koster, R. D., Mahanama, S., Livneh, B., Lettenmaier, D. P., and Reichle, R. H.: Skill in streamflow forecasts derived

1035   from large-scale estimates of soil moisture and snow, Nature Geoscience, 3, 613-616, 2010.

1036   Malbéteau, Y., Merlin, O., Molero, B., Rüdiger, C., and Bacon, S.: DisPATCh as a tool to evaluate coarse-scale remotely

1037   sensed soil moisture using localized in situ measurements: Application to SMOS and AMSR-E data in Southeastern

1038   Australia, Int J Appl Earth Obs, 45, 221-234, https://doi.org/10.1016/j.jag.2015.10.002, 2016.

1039   Meesters, A. G. C. A., De Jeu, R. A. M., and Owe, M.: Analytical derivation of the vegetation optical depth from the

1040   microwave polarization difference index, IEEE Geosci. Remote Sens. Lett., 2, 121-123, 2005.

1041   Mendoza, P. A., Mizukami, N., Ikeda, K., Clark, M. P., Gutmann, E. D., Arnold, J. R., . . . Rajagopalan, B.: Effects of

1042   different regional climate model resolution and forcing scales on projected hydrologic changes, J. Hydrol., 541, 1003-1019,

1043   https://doi.org/10.1016/j.jhydrol.2016.08.010, 2016.

1044   Meng, X. J., Mao, K. B. A., Meng, F., Shi, J. C., Zeng, J. Y., Shen, X. Y., . . . Guo, Z. H.: A fine-resolution soil moisture

1045   dataset for China in 2002-2018, Earth System Science Data, 13, 3239-3261, 10.5194/essd-13-3239-2021, 2021.

1046   Merlin, O., Al Bitar, A., Walker, J. P., and Kerr, Y.: An improved algorithm for disaggregating microwave-derived

1047   soil moisture based on red, near-infrared and thermal-infrared data, Remote Sens. Environ., 114, 2305-2316,

1048   https://doi.org/10.1016/j.rse.2010.05.007, 2010.

1049   Merlin, O., Walker, J. P., Chehbouni, A., and Kerr, Y.: Towards deterministic downscaling of SMOS soil moisture

1050   using MODIS derived soil evaporative efficiency, Remote Sens. Environ., 112, 3935-3946,

1051   http://doi.org/10.1016/j.se.2008.06.012, 2008.

1052   Merlin, O., Chehbouni, A. G., Kerr, Y. H., Njoku, E. G., and Entekhabi, D.: A combined modeling and

1053   multipectral/multiresolution remote sensing approach for disaggregation of surface soil moisture: Application to SMOS

1054   configuration, IEEE Trans. Geosci. Remote Sens, 43, 2036-2050, http://doi.org/10.1109/TGRS.2005.853192, 2005.

1055   Merlin, O., Escorihuela, M. J., Mayoral, M. A., Hagolle, O., Al Bitar, A., and Kerr, Y.: Self-calibrated evaporation-

1056   based disaggregation of SMOS soil moisture: An evaluation study at 3 km and 100 m resolution in Catalunya, Spain,

1057   Remote Sens. Environ., 130, 25-38, 10.1016/j.rse.2012.11.008, 2013.

1058   Merlin, O., Malbeteau, Y., Notfi, Y., Bacon, S., Er-Raki, S., Khabba, S., and Jarlan, L.: Performance Metrics for Soil

1059   Moisture Downscaling Methods: Application to DISPATCH Data in Central Morocco, Remote Sens., 7, 3783-3807,

1060   http://doi.org/10.3390/rs70403783, 2015.

1061   Molero, B., Merlin, O., Malbéteau, Y., Al Bitar, A., Cabot, F., Stefan, V., . . . Jackson, T. J.: SMOS disaggregated soil

1062   moisture product at 1km resolution: Processor overview and first validation results, Remote Sens. Environ., 180, 361-376,

1063   http://doi.org/10.1016/j.rse.2016.02.045, 2016.

1064   Montaldo, N., Albertson, J. D., Mancini, M., and Kiely, G.: Robust simulation of root zone soil moisture with

1065   assimilation of surface soil moisture data, Water Resour Res, 37, 2889-2900, 10.1029/2000WR000209, 2001.

O'Neill, P. E., Bindlish, R., Chan, S., Chaubell, J., Colliander, A., Njoku, E., and Jackson, T.: SMAP Algorithm Theoretical Basis Document: Level 2 & 3 Soil Moisture (Passive) Data Products, Revision G., Jet Propulsion Laboratory, Pasadena, CA, 2021.

Owe, M., de Jeu, R., and Walker, J.: A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index, IEEE Trans. Geosci. Remote Sens, 39, 1643-1654, 2001.

Pan, H., Chen, Z., Wit, A. D., and Ren, J.: Joint Assimilation of Leaf Area Index and Soil Moisture from Sentinel-1 and Sentinel-2 Data into the WOFOST Model for Winter Wheat Yield Estimation, Sensors (Basel, Switzerland), 19, 2019.

Peng, J., Loew, A., Zhang, S. Q., Wang, J., and Niesel, J.: Spatial downscaling of satellite soil moisture data using a vegetation temperature condition index, IEEE Trans. Geosci. Remote Sens, 54, 558-566, http://doi.org/10.1109/TGRS.2015.2462074, 2016.

Piles, M., Entekhabi, D., and Camps, A.: A Change Detection Algorithm for Retrieving High-Resolution Soil Moisture From SMAP Radar and Radiometer Observations, IEEE Trans. Geosci. Remote Sens, 47, 4125-4131, 10.1109/TGRS.2009.2022088, 2009.

Sabaghy, S., Walker, J. P., Renzullo, L. J., Akbar, R., Chan, S., Chaubell, J., . . . Yueh, S.: Comprehensive analysis of alternative downscaled soil moisture products, Remote Sens. Environ., 239, 111586, https://doi.org/10.1016/j.rse.2019.111586, 2020.

Sanchez-Ruiz, S., Piles, M., Sanchez, N., Martinez-Fernandez, J., Vall-Ilossera, M., and Camps, A.: Combining SMOS with visible and near/shortwave/thermal infrared satellite data for high resolution soil moisture estimates, J. Hydrol., 516, 273-283, 10.1016/j.jhydrol.2013.12.047, 2014.

Scaini, A., Sanchez, N., Vicente-Serrano, S. M., and Martinez-Fernandez, J.: SMOS-derived soil moisture anomalies and drought indices: a comparative analysis using in situ measurements, Hydrological Processes, 29, 373-383, 10.1002/hyp.10150, 2015.

Song, P. and Zhang, Y.: An improved non-linear inter-calibration method on different radiometers for enhancing coverage of daily LST estimates in low latitudes, Remote Sens. Environ., 264, 112626, https://doi.org/10.1016/j.rse.2021.112626, 2021a.

Song, P. and Zhang, Y.: Daily all weather surface soil moisture data set with 1 km resolution in China (2003-2019), National Tibetan Plateau Data Center [dataset], 10.11888/Hydro.tpdc.271762, 2021b.

Song, P., Huang, J., and Mansaray, L. R.: An improved surface soil moisture downscaling approach over cloudy areas based on geographically weighted regression, Agr Forest Meteorol, 275, 146-158, 10.1016/j.agrformet.2019.05.022, 2019a.

1095   Song, P., Zhang, Y., and Tian, J.: Improving Surface Soil Moisture Estimates in Humid Regions by an Enhanced

1096   Remote Sensing Technique, Geophys Res Lett, 48, e2020GL091459, https://doi.org/10.1029/2020GL091459, 2021.

1097   Song, P., Mansaray, L. R., Huang, J., and Huang, W.: Mapping paddy rice agriculture over China using AMSR-E

1098   time series data, Isprs J Photogramm, 144, 469-482, 10.1016/j.isprsjprs.2018.08.015, 2018.

1099   Song, P., Huang, J., Mansaray, L. R., Wen, H., Wu, H., Liu, Z., and Wang, X.: An Improved Soil Moisture Retrieval

1100   Algorithm Based on the Land Parameter Retrieval Model for Water-Land Mixed Pixels Using AMSR-E Data, IEEE Trans.

1101   Geosci. Remote Sens, 1-15, 10.1109/TGRS.2019.2915346, 2019b.

1102   Stefan, V. G., Merlin, O., Escorihuela, M.-J., Molero, B., and Er-Raki, S.: Temporal Calibration of an Evaporation-

1103   Based Spatial Disaggregation Method of SMOS Soil Moisture Data, Remote Sens., 12, 1671, 2020.

1104   Sui, D. Z.: Tobler's First Law of Geography: A Big Idea for a Small World?, Annals of the Association of American

1105   Geographers, 94, 269-277, https://doi.org/10.1111/j.1467-8306.2004.09402003.x, 2004.

1106   Ulaby, F. T. and Wilson, E. A.: Microwave Attenuation Properties of Vegetation Canopies, IEEE Trans. Geosci.

1107   Remote Sens, GE-23, 746-753, 10.1109/TGRS.1985.289393, 1985.

1108   Vergopolan, N., Xiong, S. T., Estes, L., Wanders, N., Chaney, N. W., Wood, E. F., . . . Sheffield, J.: Field-scale soil

1109   moisture bridges the spatial-scale gap between drought monitoring and agricultural yields, Hydrol. Earth Syst. Sci., 25,

1110   1827-1847, 2021.

1111   Verstraeten, W. W., Veroustraete, F., van der Sande, C. J., Grootaers, I., and Feyen, J.: Soil moisture retrieval using

1112   thermal inertia, determined with visible and thermal spaceborne data, validated for European forests, Remote Sens.

1113   Environ., 101, 299-314, 2006.

1114   Wang, K. and Liang, S.: Evaluation of ASTER and MODIS land surface temperature and emissivity products using

1115   long-term surface longwave radiation observations at SURFRAD sites, Remote Sens. Environ., 113, 1556-1565,

1116   https://doi.org/10.1016/j.rse.2009.03.009, 2009.

1117   Wang, L. and Qu, J. J.: NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture

1118   with satellite remote sensing, Geophys Res Lett, 34, L20405, 10.1029/2007GL031021, 2007.

1119   Wei, Z., Meng, Y., Zhang, W., Peng, J., and Meng, L.: Downscaling SMAP soil moisture estimation with gradient

1120   boosting decision tree regression over the Tibetan Plateau, Remote Sens. Environ., 225, 30-44, 2019.

1121   Wu, D., Liang, H., Cao, T., Yang, D., Zhou, W., and Wu, X.: Construction of operation monitoring system of automatic

1122   soil moisture observation network in China, Meteorological Science and Technology, 42, 278-282, 2014

1123   Yang, G., Sun, W. W., Shen, H. F., Meng, X. C., and Li, J. L.: An Integrated Method for Reconstructing Daily MODIS

1124   Land Surface Temperature Data, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 12, 1026-1040, 2019.

1125   Yao, P., Lu, H., Shi, J., Zhao, T., Yang, K., Cosh, M. H., . . . Entekhabi, D.: A long term global daily soil moisture

1126   dataset derived from AMSR-E and AMSR2 (2002–2019), Scientific Data, 8, 143, 10.1038/s41597-021-00925-8, 2021.

1127   Zeng, Y., Feng, Z., and Xiang, N.: Assessment of soil moisture using Landsat ETM+ temperature/vegetation index in

1128   semiarid environment, IEEE International Geoscience & Remote Sensing Symposium, Piscataway NJ, 2004, 4306-4309

1129   vol.4306, 10.1109/IGARSS.2004.1370089,

1130   Zhang, J., Zhou, Z., Yao, F., Yang, L., and Hao, C.: Validating the Modified Perpendicular Drought Index in the North

1131   China Region Using In Situ Soil Moisture Measurement, IEEE Geoscience & Remote Sensing Letters, 12, 542-546, 2014.

1132   Zhang, Y., Kong, D., Gan, R., Chiew, F. H. S., Mcvicar, T. R., Zhang, Q., and Yang, Y.: Coupled estimation of 500 m

1133   and 8-day resolution global evapotranspiration and gross primary production in 2002-2017, Remote Sens. Environ., 222,

1134   165-182, 2019.

1135   Zhang, Y. Q., Chiew, F. H. S., Liu, C. M., Tang, Q. H., Xia, J., Tian, J., . . . Li, C. C.: Can Remotely Sensed Actual

1136   Evapotranspiration Facilitate Hydrological Prediction in Ungauged Regions Without Runoff Calibration?, Water Resour

1137   Res, 56, 2020.

1138   Zheng, J. Y., Lu, H. S., Crow, W. T., Zhao, T. J., Merlin, O., Rodriguez-Fernandez, N., . . . Gou, Q. Q.: Soil moisture

1139   downscaling using multiple modes of the DISPATCH algorithm in a semi-humid/humid region, Int J Appl Earth Obs, 104,

1140   10.1016/j.jag.2021.102530, 2021.

1141   Zhou, S., Williams, A. P., Lintner, B., Berg, A. M., and Gentine, P.: Soil moisture–atmosphere feedbacks mitigate

1142   declining water availability in drylands, Nature Climate Change, 11, 2021.

1143   Zhu, Z. and Shi, C.: Simulation and Evaluation of CLDAS and GLDAS Soil Moisture Data in China (written in

1144   Chinese), Science Technology and Engineering, 32, 138-144, 2014.

1145