1 # A 1-km daily surface soil moisture dataset of enhanced coverage

2 ## under all-weather conditions over China in 2003-2019

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

5 [1] Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences

6 and Natural Resources Research, The Chinese Academy of Sciences, Beijing 100101, China

7 [2] State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing

8 100081, China

9 [3] National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China

10 [4] State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese

11 Academy of Sciences. Beijing 100101, China

12

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

14

15

16

17

18

19

20

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.

48 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

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

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

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

154 2 is operating onboard the Global Change Observation Mission1-Water (GCOM-W1)

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

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

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

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

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

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

161 developed by the University of Montana (UMT) (Jones et al., 2009; Du et al., 2016). A

162 recent AMSR-based night-time SSM product during 2002-2019 has been produced

163 through a neural network trained against SMAP radiometer-based descending SSM

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

165 results show that this NN-SM product is better than the JAXA and LPRM products.

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

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

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

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

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

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

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

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

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

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

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

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

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

179 product as described above.

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

181 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/

182

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

Optical remote sensing datasets provide finer spatial texture information on the daily basis for the downscaling purpose of PM SSM. Such data that can be used as inputs of our SSM product processing line are mainly provided by the Moderate-resolution Imaging Spectroradiometer (MODIS) onboard the Terra and Aqua satellites. Specifically, they involve the 1-km daily night-time Aqua MODIS LST product (MYD21A1N.v061) and the 500-m daily "Bidirectional Reflectance Distribution Function (BRDF)" - Adjusted Reflectance dataset (MCD43A4.v061). MYD21A1 LST data can be recognized as a crucial proxy of land surface thermal capacity (Fang et al., 2013) and soil evaporative rate (Merlin et al., 2008). The MCD43A4 nadir reflectance product, with view angle effect corrected using the BRDF model, is capable to provide observations from visible to shortwave-infrared bands that can characterize water content variation of the bare soils as well as the vegetation canopy. Overall, the above-mentioned datasets were selected primarily because they deliver indicators (land surface thermal capacity, soil evaporative rate, or vegetation condition) that can well response to soil moisture dynamics from different aspects. Prior to being employed for SSM downscaling, conventional pre-processing procedure of pixel quality check was applied for both optical products by screening out pixels not classed as "good quality", according to the 8-bit "Quality Assessment (QA)" field of each spectral band. Moreover, to normalize their natively different spatial resolutions, all MCD43A4 based reflectance values at the 500-m pixel level were upscaled to the sinusoidally projected MODIS 1-km grids using their spatial averages.

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

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

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

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

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

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

211   DEM data.

212   2.1.3 Study area and validation data

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

214   evaluate the SSM downscaling performances for different geographic regions (see

215   Section 3.3), we divided the country further into six different geographic-climate

216   regions using elevation, precipitation, hydrogeology, vegetation type, and topography.

217   The six regions include the Northeast Monsoon (NEM) region, the Northwest Arid

218   (NWA) region, the Qinghai–Tibet Plateau (QTP) region, the North China Monsoon

219   (NCM) region, the South China Monsoon (SCM) region, and the Southwest Humid

220   (SWH) region. The detailed delimitation principle of these geographic-climate regions

221   was originally described in Meng et al. (2021). The geographic zoning map is shown

222   in Fig. 1, while the corresponding shapefile boundary files can be accessed from the

223   Resource and Environment Science and Data Center of the Chinese Academy of

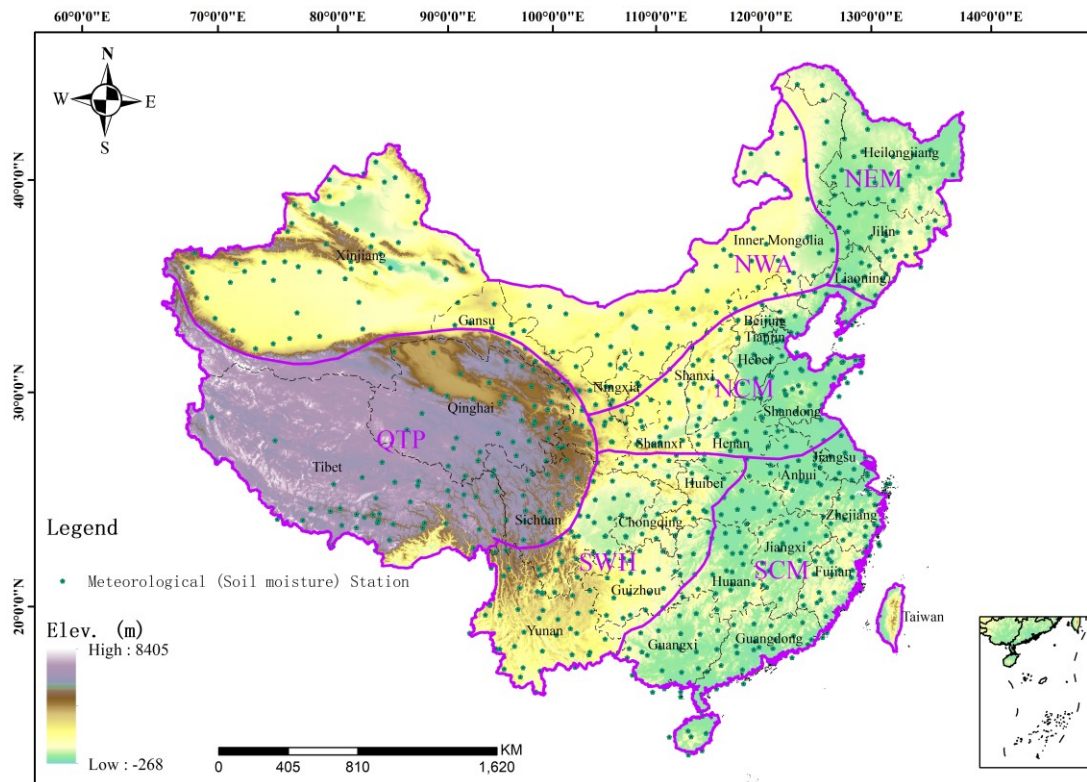224   Sciences (http://www.resdc.cn/, last access: May 22, 2021).

225

Fig. 1 The geographic zoning map of China (delineated using the purple color) superposed with

topographic information, as well as general locations for the 756 basic meteorological stations

(http://data.cma.cn/, last access: January 20, 2021) that provide partial benchmark measurements for

SSM and LST validation in this study.

We utilized ground soil moisture measurements for validating the downscaled

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

262 of its notably creditable performance (see Fig. A1 in Appendix A). The latest version

263 of this dataset (SPL3SMP, Version 8) contains soil moisture retrievals based on

264 different algorithms including the dual channel algorithm and the single channel

265 algorithm. In this study we only extracted SSM estimates derived with the dual channel

266 algorithm because this algorithm was reported to outperform the single channel

267 algorithm over some agricultural cropland core validation sites (O'neill et al., 2021).

268 2.1.4 Ancillary SSM products for comparison

269 In order to comprehensively demonstrate the validation performance of our

270 proposed SSM product, there is necessity to make an inter-comparison against similar

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

product (Section 2.1.3).

2.2   Methodology

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

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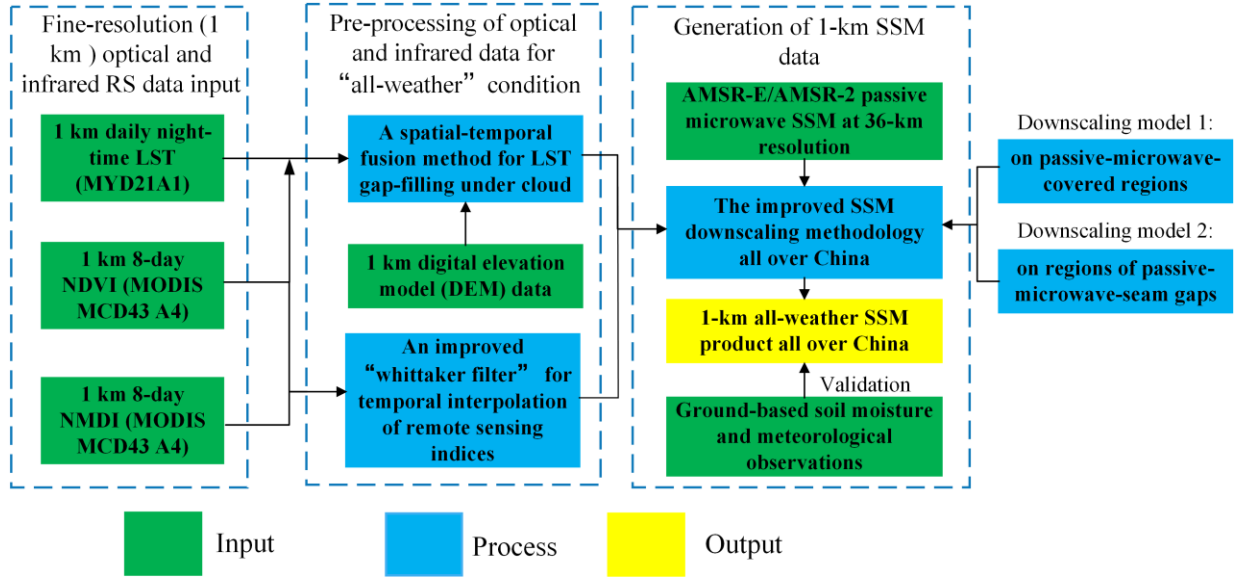
this section.

Fig. 2 The overall methodological framework of this study.

2.2.1 Reconstruction of thermal-infrared LST and remote sensing (vegetation) indices under cloud

Reconstruction of missing pixels under cloud in the optical remote sensing input datasets is the prerequisite for achieving the "all-weather" property of the final downscaled SSM output. For reconstructing thermal-infrared LST, we adopted the cloud gap-filling method as proposed by our previous study (Song et al., 2019a). This method, also referred to as a typical "spatio-temporal data fusion" (STDF) method (Dowling et al., 2021), was built using clear-sky LST observations of spatially neighboring pixels observed at proximal dates, with concurrent NDVI and DEM also employed as additional data inputs. The STDF method can be expressed as follows:

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

Where the superscript "*" indicates that this variable has been normalized to the range 0 to 1.0 (Song et al., 2019a), based on the maximum and minimum values of that variable found across China (excluding invalid values representing states of snow, ice, and water bodies). Parameters a, b, c, and d are coefficients fitted between all pixels with clear-sky LST estimates on a specific date t_1 ($LST^*_{t_1}$) and their counterparts on

16

316   one proximal date, t_0 (LST^*_{t0}). $NDVI^*_{t1}$ indicates the corresponding (normalized) NDVI

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

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

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

320   all neighboring dates of t_1 meeting the condition $|t_0 - t_1| <= 30$ (from the nearest date to

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

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

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

324   exceeded 0.99.

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

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

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

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

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

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

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

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

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

334   discussed in Section 4.2.

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

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

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

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

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

340   series-filtering-like algorithm.

341 2.2.2   Improved downscaling technique of SSM based on fusion of PM and

342 optical/infrared data

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

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

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

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

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

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

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

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

351
$$SSM = \frac{a \times \ln(1 - SEE)}{1 - b \times NMDI} + c \tag{2}$$

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

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

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

355 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).

356 $R_{infr,860nm}$, $R_{infr,1600nm}$, and $R_{infr,2100nm}$ represent land surface reflectance signals

357 derived   from   three   different   MODIS-MCD43A4   based   near-infrared/shortwave-

358 infrared bands, with their wavelengths centering at 860 nm, 1600 nm, and 2100 nm

359 respectively. The parameters *a, b,* and *c* are empirical coefficients that represent

360 background information of local soil texture and vegetation types. In Song et al. (2021),

361 these coefficients have been fitted and calibrated based on multi-temporal observations

362 at the PM pixel scale. In our current study, however, we have discovered that coupling

363 of multiphase observations at both the spatial and the temporal dimensions can lead to

364 more optimal solution of the coefficients, as they can produce downscaled SSM images

365 with notably declined effect of 'mosaic' against the original PM 36-km pixels.

366 Therefore, the modified optimal cost function χ^2 for deriving these coefficients is re-

367 defined as follows:

368
$$\chi^2 = \sum_{d=-dl}^{dl} \sum_{i=0}^{N=ws \times ws} w_i \times (SSM_{ob,i,d} - SSM_{mod,i,d})^2 \tag{3}$$

369 Through the cost function, the spatial extent of each 36-km pixel P_0 on any arbitrary

370 date D_0 obtains a unique set of coefficients. As shown by Equation (3), all pixels were

371 exploited within the spatial square window (with its side length equal to ws) centered

372 at P_0 ranging from $-dl$-th day to dl-th day relative to the date of D_0. To determine the

373 optimum values for dl and ws, we have tested each member in the collection of [3, 5, 7,

374 9, 11, 13] for both parameters. Evaluation against in-situ data indicates that the

375 optimum dl and ws are 5 and 7, respectively (results are similar to what is shown in

376 Section 3.2, but not presented here). SSM_{ob} and SSM_{mod} denote the AMSR NN-SM 36-

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

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

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

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

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

382
$$w_i = [1-(\frac{dis_i}{b})^2]^2, dis_i < b$$
$$w_i = 0, dis_i >= b \tag{4}$$

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

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

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

386 by Brunsdon et al. (1996).

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

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

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

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

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

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

393 described below:

394 $SSM_{1\text{-}km} = SSM_{36km} + (\frac{\partial SSM}{\partial SEE})_{36km} \times (SSE_{1km} - <SSE>_{36km}) + 0.5 \times (\frac{\partial^2 SSM}{\partial SEE^2}) \times (SSE_{1km} - $

395 $<SSE>_{36km})^2 + (\frac{\partial SSM}{\partial NMDI})_{36km} \times (NMDI_{1km} - <NMDI>_{36km}) + 0.5 \times (\frac{\partial^2 SSM}{\partial NMDI^2}) \times$

396 $(NMDI_{1km} - <NMDI>_{36km})^2$ (5)

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

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

399 $\frac{\partial SSM}{\partial SEE}$ $(\frac{\partial^2 SSM}{\partial SEE^2})$ and $\frac{\partial SSM}{\partial NMDI}$ $(\frac{\partial^2 SSM}{\partial NMDI^2})$ respectively denoting the first-(second-) order

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

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

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

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

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

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

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

407 $\Delta SSM_{36km,t0}$ is a component for correcting inter-day bias, with the following expression:

408
$$\Delta SSM_{36km,t0} = SSM\left(SEE_{36km,t0},\ NMDI_{36km,t0}\right) - \\ 0.5 \times \left(SSM\left(SEE_{36km,t0-1},\ NMDI_{36km,t0-1}\right) + SSM\left(SEE_{36km,t0+1},\ NMDI_{36km,t0+1}\right)\right) \tag{6}$$

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

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

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

412 1-km. If all $SSM_{36\text{-}km}$ during the three consecutive days (t_0-1, t_0, and t_0+1) are missing

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

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

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

416 ## 2.2.3 Evaluation metrics

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

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

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

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

421 more commonly in previous studies. This is because both ground observations and

422 other benchmark data (i.e. SMAP radiometer-based SSM) may also present

423 measurement uncertainties in practice. For SSM evaluation, the unbiased RMSD, or

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

425 when validated against ground soil moisture measurements. This can better investigate

426 the time series similarity between satellite and in situ datasets by eliminating the

427 systematic bias caused by spatial scale mismatch between them.

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

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

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

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

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

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

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

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

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

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

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

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

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

3. Results

441

3.1 Evaluation on reconstructed thermal-infrared LST under

442

443   cloud

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

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

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

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

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

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

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

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

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

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

454   2.1.3) where 0-cm in-situ top-ground temperature measurements were available, we

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

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

457   these stations, with the results revealed in Fig. 3. It is manifested through Fig. 3-(a) and

458   -(b) that very close performances have been achieved between the clear-sky and the

459   cloudy scenarios, especially considering their almost equally high validating

460   correlations between 0.94-0.96. For each independent station, we calculated the

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

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

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

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

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

466   is about 1.9 K. All above results have indicated that the uncertainty of our night-time

467   LST reconstruction algorithm proposed for cloudy conditions is not very significant.

468 The corresponsive uncertainty that could be propagated to downscaled SSM in this
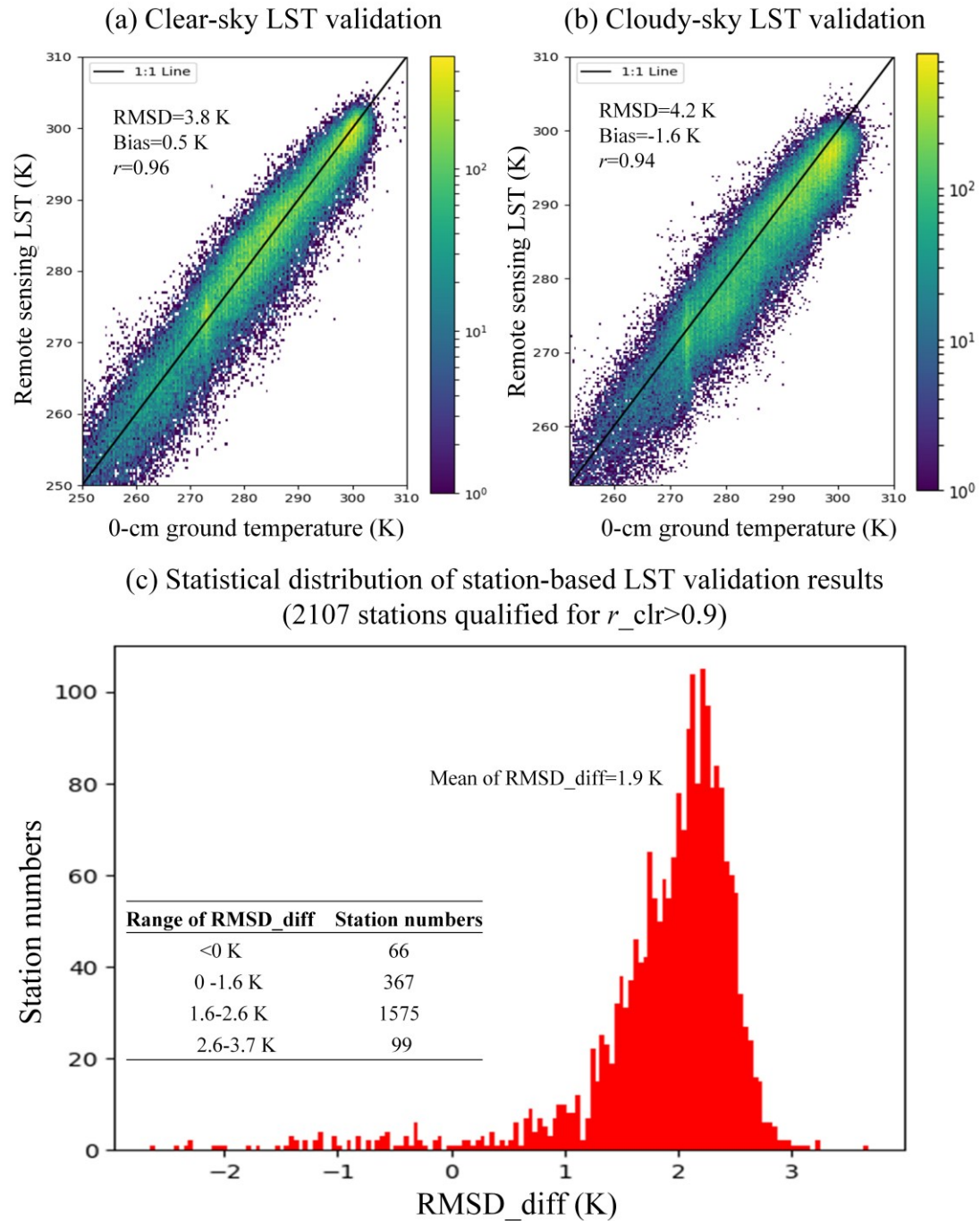
469 stage is analyzed below in Section 3.2.

(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

Mean of RMSD_diff=1.9 K

470

471 Fig. 3 validation results of the cloud gap-filled LST in China. (a) Density plot of thermal infrared

472 LST under clear-sky condition compared to the 0-cm ground temperature measurements for all

473 stations. (b) Same to (a) but for thermal infrared LST under cloudy conditions. (c) Statistical
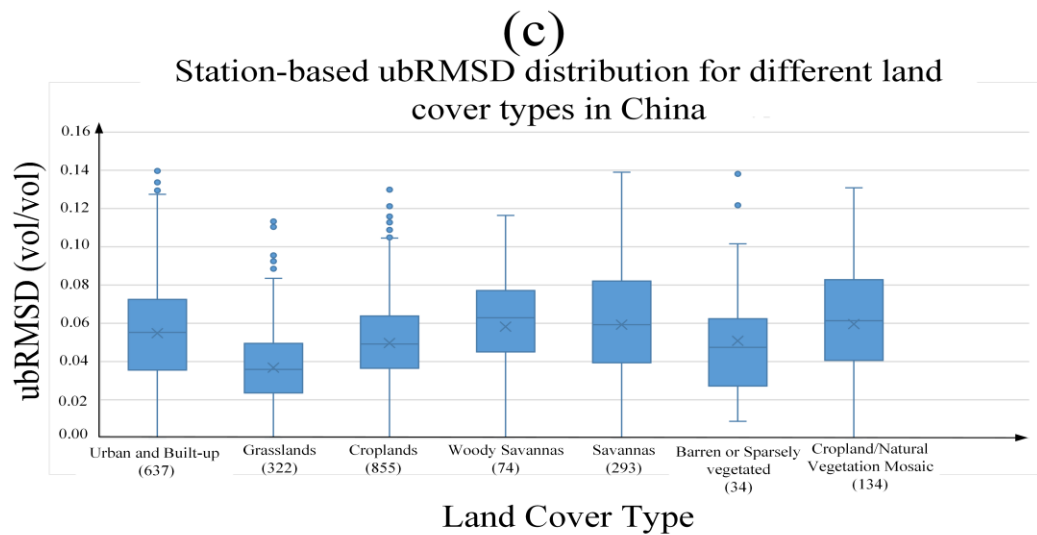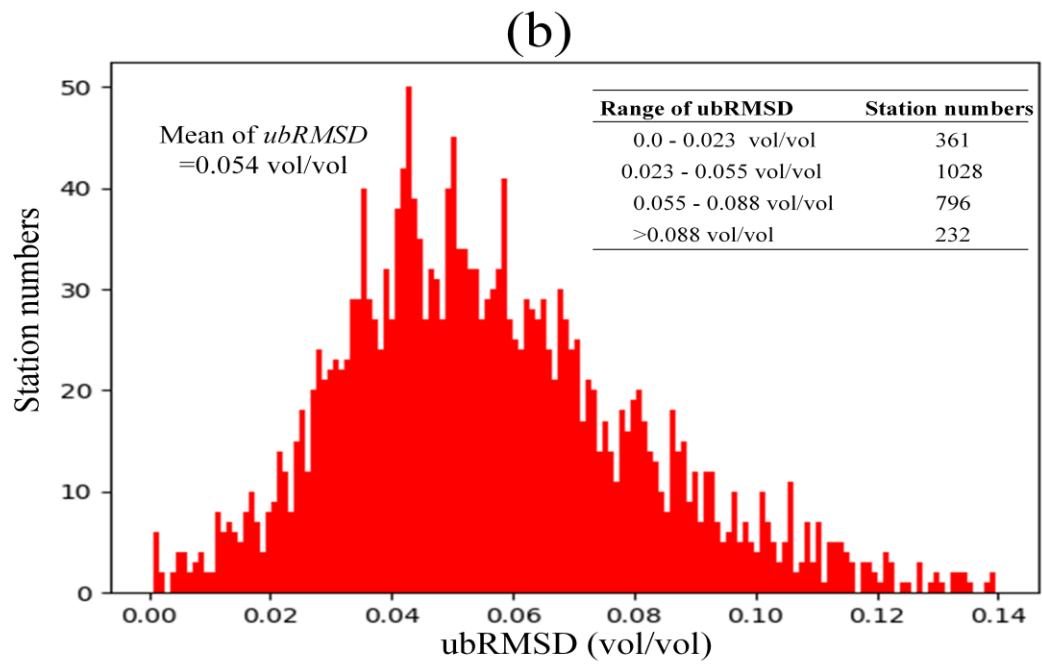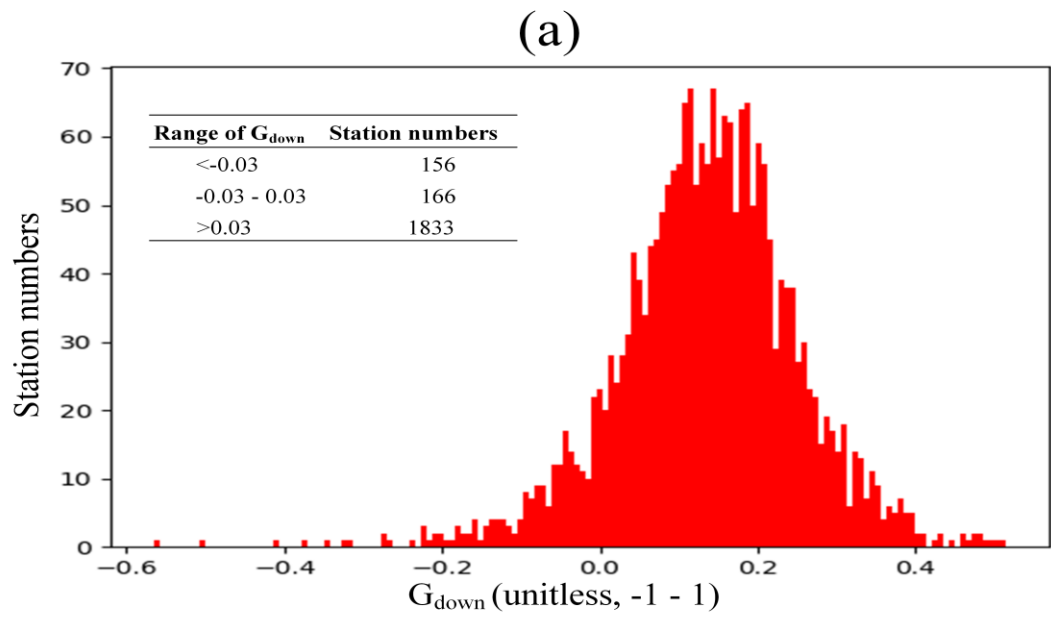
474   distribution of difference between RMSD of clear-sky LST and RMSD of gap-filled LST under cloudy

475   condition with regard to different meteorological stations over the study region.

476

477 3.2 Evaluation on the final 1-km SSM product

478   The overall validation results of the finally downscaled 1-km SSM product against

479 ground soil moisture data is demonstrated in Fig. 4. Fig. 4-(a) shows that about 85%

480 (N: 1833) of the total 2154 stations (the remaining 263 stations are located in pixels

481 with no effective PM observations and are thus removed) have obtained significantly

482 positive downscaling gains (G_{down}>0.03). This hints that the 1-km SSM product can

483 better capture the dynamic behaviors of local ground soil moisture data than the original

484 36-km PM NN-SM data, revealing higher spatial representativeness of the downscaled

485 SSM data product over the country. According to Fig. 4-(b), the mean ubRMSD of all

486 stations is about 0.054 vol/vol, while 90% of those stations have the number lower than

487 0.088 vol/vol. In addition, we made another analysis concerning the possible influence

488 of land cover types on SSM downscaling performance in Fig. 4-(c). The spatial

489 information of land cover types was derived from the MODIS MCD12Q1

490 (10.5067/MODIS/MCD12Q1.006) IGBP-based land use image in 2019. For stations

491 that experienced land use change throughout the years of the study period, the ubRMSD

492 is only reported for data in the year of 2019. Clearly, better accuracies are observed

493 mainly in grassland, cropland and bare soil surface, whilst relatively poorer

494 performances (with averages of ubRMSD higher than 0.06 vol/vol) are seen in urban

495 regions, (woody) savanna, and crop-to-natural-vegetation mosaic areas. Such a relative

496   performance across land covers is logical because all the land cover types with their

497   average ubRMSD higher than 0.06 vol/vol are characterized by lower hydrologic

498   homogeneity in terms of their definition, e.g. savanna, which is a mixture of grass and

499   tall trees, and urban areas, which are composed of impervious 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)

Urban and Built-up (637)   Grasslands (322)   Croplands (855)   Woody Savannas (74)   Savannas (293)   Barren or Sparsely vegetated (34)   Cropland/Natural Vegetation Mosaic (134)

Land Cover Type

500

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

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

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

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

505   In Fig. 5, we compared time series of regionally aggregated SSM from our

506   developed 1-km SSM product to that from the SMAP 36-km descending SSM, for each

507   of the six different geographic-climate regions (as shown in Fig. 1) from 2016 to 2019.

508   Via this effort, we mainly aim to reveal the consistency degree on reflecting soil

509   moisture temporal dynamics at different geographical settings between the two SSM

510   products. This also provides another view to evaluate the reliability of our developed

511   product. Because the SMAP radiometer has a slightly longer revisit cycle (~2-3 days)

512   than AMSR-2, the time series data are also aggregated and averaged at the temporal

513   dimension, with a displayed revisit cycle equal to three days. Overall, the time series

514   data correlate well with each other for all six regions. The relatively lower RMSDs

515   (<0.02 vol/vol) are found in regions with comparatively sparser vegetation covers

516   including the NWA region, the QTP region, and the NCM region. For other three dense-

517   vegetation regions, the performances of our developed product are slightly poorer. This

518   is especially the case for the SCM region, with a lower r-value of 0.84. The reason can

519   be attributed to the enlarged difference on penetration depth into the soil layers between

520   L-band (SMAP) and C-/X-/K- band (AMSR-2) emissions under dense vegetation

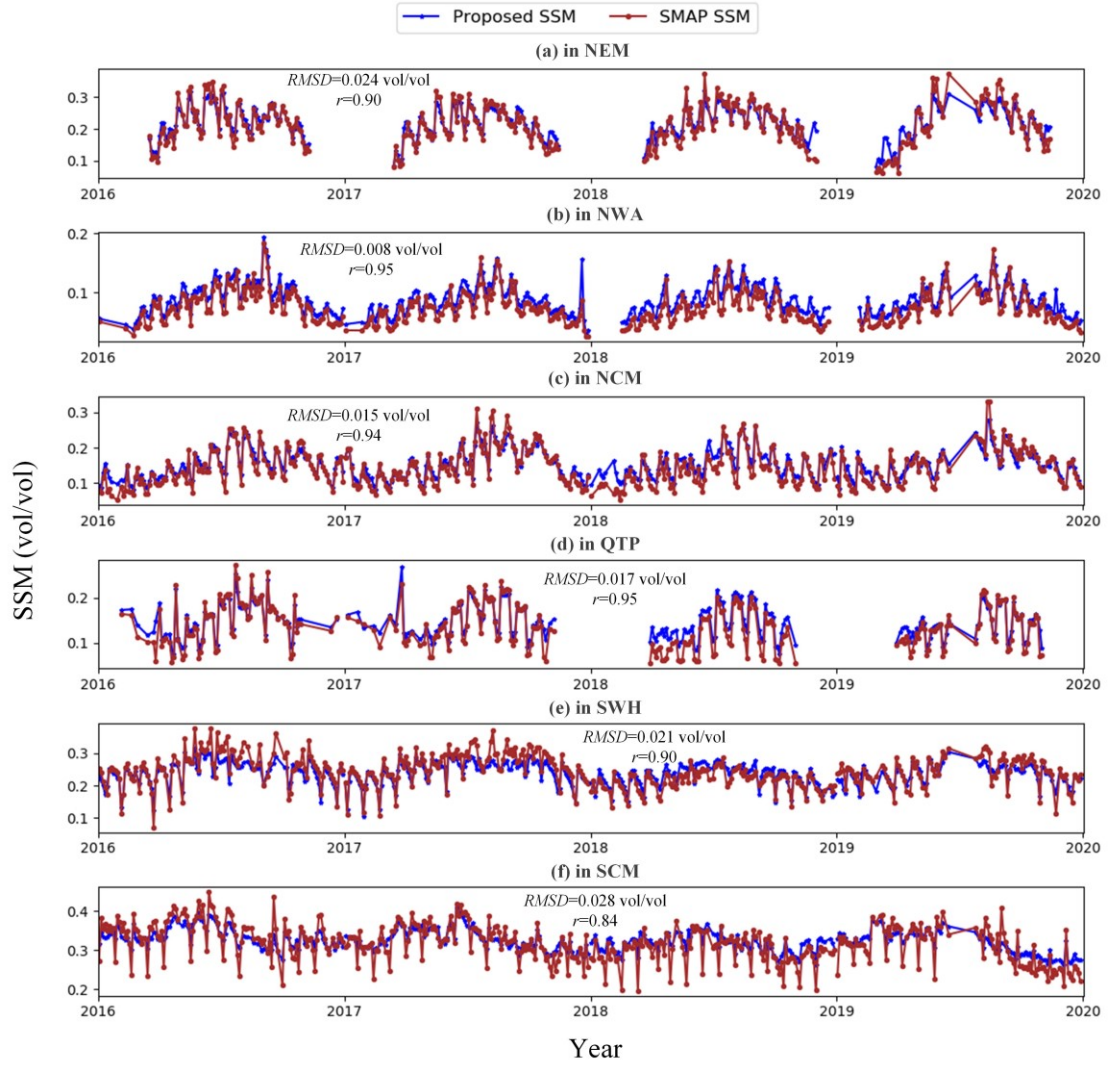521   covers (Ulaby and Wilson, 1985).

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 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 26 to May 31, 2018 in Fig. 6-(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 product on one single

532 date, even though the former was generated based on multi-date images. Both maps

533 show similar spatial texture depicting the relatively dry climate in northwestern China

534 compared with the humid climate in the Middle-lower Yangtze River Plain.

535 Nevertheless, there also exist cases where the details in texture differ prominently, like

536 that in the far northeastern end of the country.

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

538 as proposed in Section 2.1.3 and Section 2.1.4, is demonstrated in Fig. 6-(c), -(d), -(e),

539 and -(f). Fig. 6-(c) and -(d) show the RMSD maps of the two respective products against

540 SMAP radiometer-based SSM estimates at the 36-km pixel scale. For both products it

541 is manifested that compared with the lower averaged RMSD of 0.04 vol/vol in the

542 NWA region, the uncertainty can increase (shown in yellow) in the densely vegetated

543 NEM and the SCM regions, with averaged RMSDs of 0.07-0.08 vol/vol. However, our

544 developed product has noticeably lower RMSD (0.05 vol/vol) than the

545 SPL2SMAP_S_V3 data (0.07-0.09 vol/vol) in the SWH and part of the QTP regions.

546 Considering their relatively higher elevations, it may be roughly drawn that our

547 downscaled SSM product is more reliable than that downscaled based on active-passive

548 microwave combined datasets in areas with increased topographic effects. Fig. 6-(e)

549 shows that the currently developed SSM product obtained a 0.078 vol/vol ubRMSD

550 and a correlation coefficient of 0.55 against the in-situ soil moisture measurements. It

551 converges more apparently to the 1:1 line when compared with validation result of the

552 SPL2SMAP_S_V3 dataset in Fig. 6-(f). As with the area of China, therefore, the

553   currently developed product is generally superior to the global SMAP/Sentinel

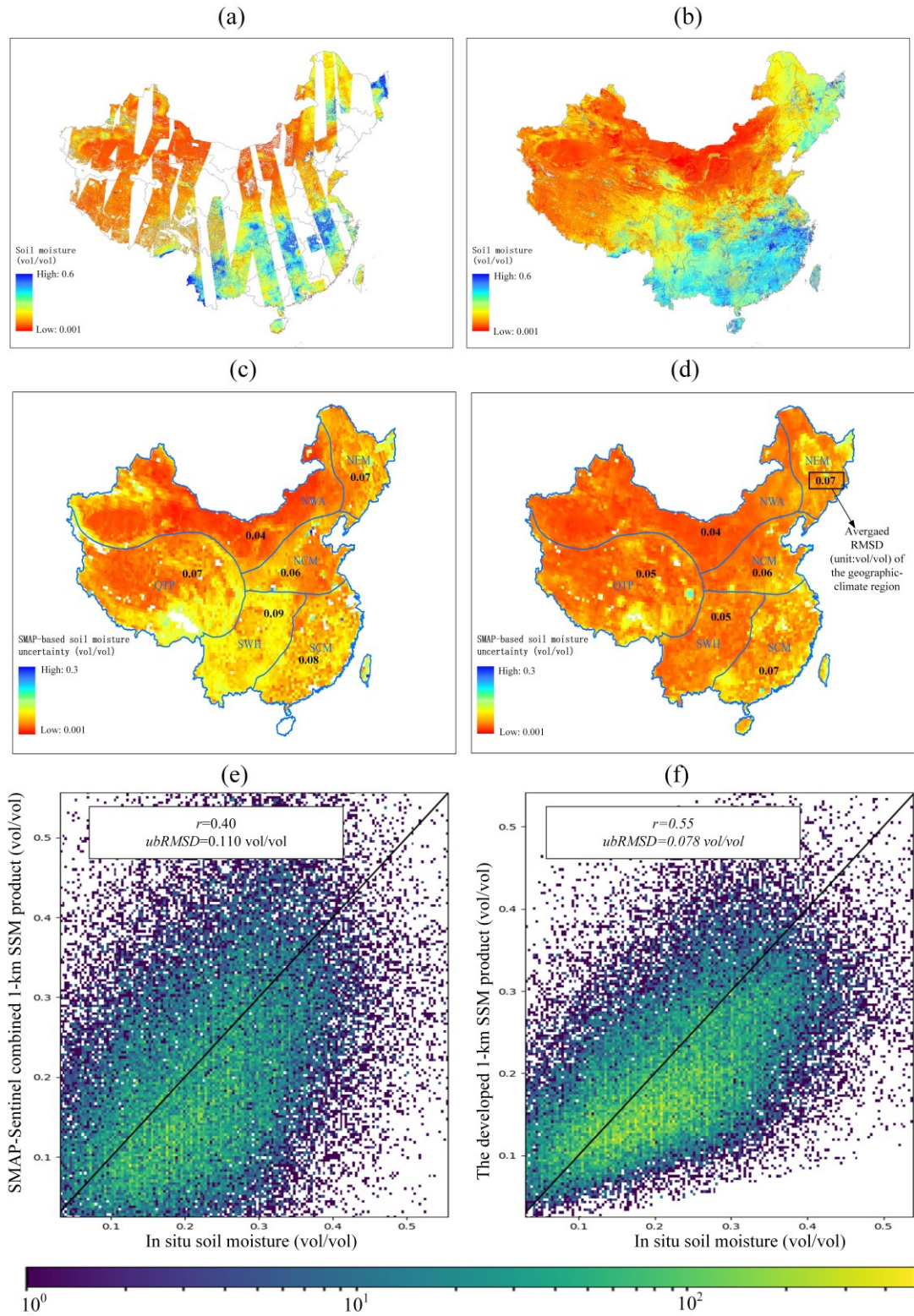554   combined SSM in terms of both coverage percentage and estimate accuracy.

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 May 26, 2018 to May 31, 2018. (b) The SSM image at

559   1:30 a.m. of May 29, 2018 from the currently developed product. (c) Spatial uncertainty (RMSD) map

560   of the SPL2SMAP_S_V3 product against SMAP radiometer-based SSM retrievals at the 36-km pixel

561   scale over China for years of 2017, 2018, and 2019. (d) Same to (c) but for validaiton of the currently

562   developed SSM product. The black numbers in each of the geographic-climate regions indicate

563   averaged uncertainty (RMSD, unit: vol/vol) of the region. (e) Validation results of the

564   SPL2SMAP_S_V3 product against in-situ soil moisture measurements over China for years of 2017,

565   2018, and 2019. The black solid line is the 1:1 line. (f) Same to (e) but for validaiton of the currently

566   developed SSM product.

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

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

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

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

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

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

573   2011 to June 2012, during which there are no effective observations from the PM NN-

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

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

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

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

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

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

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

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

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

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

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

585   time. The above inter-seasonal differences on data coverage are also reflected in Fig. 8

586   in another manner based on presenting the spatial distributions of number percentages

587   of available dates in each three-month period.

588
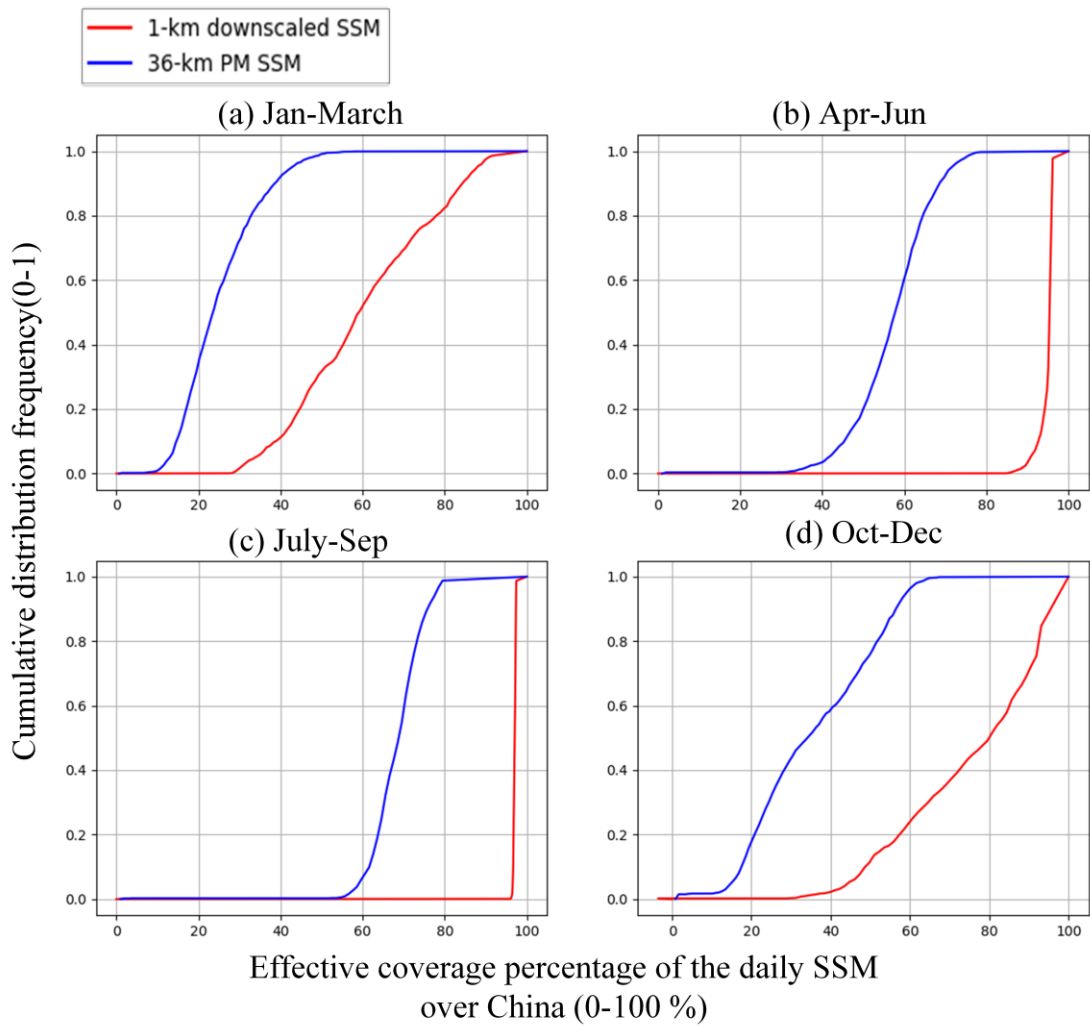
589   Fig. 7 Cumulative distribution frequency of our proposed SSM product against the original 36-km SSM

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

591   current statistics.
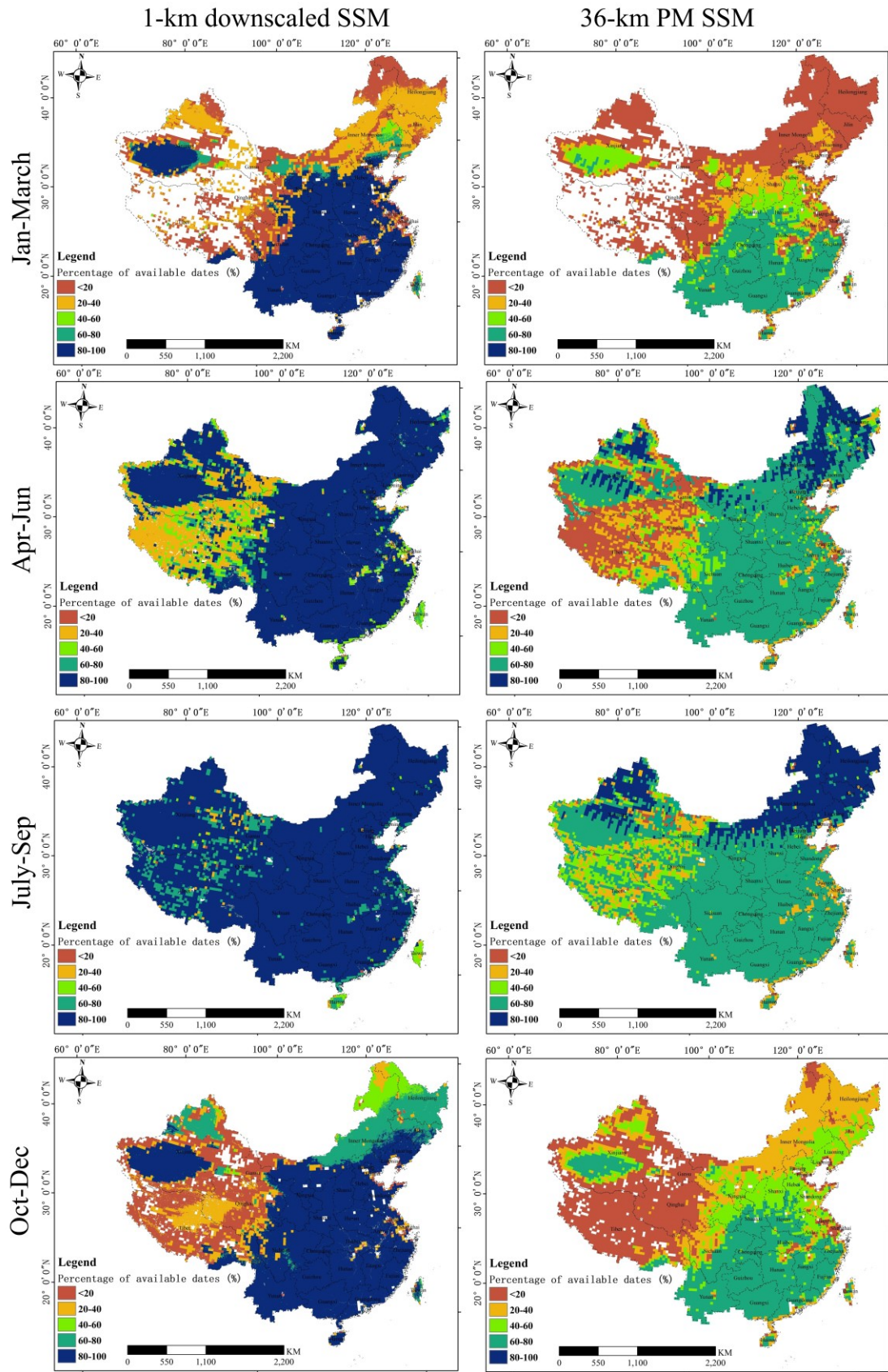
592

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

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

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

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

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

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

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

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

602   (using averages of ground measurements at all stations) of downscaled SSM in these

603   coverage-improved conditions, relative to that generated without using any coverage

604   improvement technique, in order to evaluate the propagated effect of such techniques

605   on the final product. The very limited difference for ubRMSD values (0.053 vol/vol

606   versus 0.056 vol/vol) between cloudy and clear-sky conditions suggest that the 1-km

607   SSM estimates from our final product are generally compatible between cloudy and

608   clear-sky conditions. The downscaled SSM estimated for regions of PM-seam gaps

609   have a slightly worse (but still acceptable) accuracy, considering its ubRMSD of 0.059

610   vol/vol compared to the 0.052 vol/vol ubRMSD of the PM-observed 1-km pixels. In

611   summary of Fig. 7 and Table 2, the currently developed product has achieved a

612   substantially improved spatial coverage against the original remote sensing input

613   datasets, whilst successfully preserved the SSM downscaling accuracy of the

614   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

4. Discussion

4.1 Uncertainty 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

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

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

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

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

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

636 coarser resolution. But on the other hand, only a few ($\leq$4) of such "integrated stations"

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

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

639 network of China exploited by our current study.

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

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

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

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

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

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

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

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

648 combination of multi-station ground measurements shows a global ubRMSD of 0.078

649 vol/vol for this product. Overall, the above-mentioned results can be identified as at

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

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

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

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

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

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

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

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

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

659 analysis of satellite remote sensing soil moisture. However, we should be noticed that

660 these previous studies have conducted analyses respectively at the temporal and the

661 spatial dimensions. Based on their results, the spatial analysis typically derived lower

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

663 the heterogeneity degree of remote sensing pixels can vary significantly across different

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

665 dimensions, such an r-value is expected. This is also close to the global r-value of 0.6

666 for validation of the coarse-resolution NN-SM product as reported in Yao et al. (2021).

667 ## 4.2   Uncertainty on cloud gap-filling and validations of LST

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

691 final SSM product may be even more serious than indication from the above RMSD

692 difference (i.e. 0.056-0.053 = 0.003 vol/vol), due to the portion of "clear/cloudy-

693 weather-mixed" spatial windows during the fitting process of the downscaling model.

694 In these windows, uncertainty in cloud gap-filled LST may affect accuracy of the fitted

695 model coefficients and thus deteriorate the final SSM estimates in clear-sky pixels

696 within the same window. Consequently, the above RMSD difference has been more or

697　less underestimated. Despite all of above, in our study area of China we regard the

698　STDF-filled LST as a more optimal proxy of heat flux for estimating SSM under clouds,

699　compared to the bias-adjusted LST. On the other hand, future efforts are encouraged to

700　further clarify the mechanical relationships between STDF-filled/bias-adjusted LST

701　and soil wetness degree under clouds.

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

703　2021; Yang et al., 2019) validating satellite thermal-infrared-based LST based on

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

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

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

707　motivation driving such an experimental design is the significantly intensive

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

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

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

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

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

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

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

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

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

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

718   marginal effect on our results. We have carried an extra test that can confirm this

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

720 4.3 Major novelty, unique profit, and future prospect of the

721 developed product

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

737   herein confirm that the accuracy of the product is stable and consistent among all

738   weather conditions.

739   With improvement achieved at the three dimensions, unique profit of the currently

740 developed product can be taken by subsequent studies and various industrial

741 applications. For example, the capability of this product can be investigated on

742 capturing the short-term anomaly of local hydrological signals as well as improved

743 monitoring on drought disasters, which used to be investigated mainly at a coarser

744 resolution by PM SSM (Scaini et al., 2015; Champagne et al., 2011; Albergel et al.,

745 2012). For another, taking advantage of its all-weather daily time series, the product

746 can be utilized together with precipitation data to isolate and quantify the anthropic

747 influence on regional water resources from the natural hydrological dynamics.

748 Examples of such anthropic signals include agricultural irrigation activities, as well as

749 finer-scale information on agricultural crops which was previously interpreted based on

750 PM-driven techniques (Song et al., 2018). In addition, we should realize the important

751 role of soil moisture as a constraint for accurate estimation of surface

752 evapotranspiration and runoff (Zhang et al., 2020; Zhang et al., 2019). Therefore, the

753 profit of this product can be further enhanced if coupled with land-atmosphere coupled

754 models to produce new insights into water-cycle processes of earth surface at a finer

755 spatio-temporal scale.

756   There are still some limitations on our current product to be further improved. First,

757 there may exist the 'mosaic effect' at the original PM (36-km) pixel edge. As mentioned

758 in Section 2.2.2, we have used a parameter of 'spatial square window (*ws*)' in Equation-

759 (3) to minimize this negative effect. However, it is still difficult to utterly avoid such

760 negative effect. This is a challenge for all existing SSM downscaling methods (Molero

761 et al., 2016; Stefan et al., 2020; Peng et al., 2016), especially considering the large

762 spatial scale of our study and all uncertainties discussed in Sections 4.1 and 4.2. Besides,

763 other negative influences can be imposed by potential imperfections identified from the

764 original PM product, e.g. from PM SSM retrievals in the QTP region with complicated

765 topography, melt snow or partially frozen soils that cannot been completely screened

766 out by the original product flag in winter. For these extreme conditions, the accuracy

767 of the downscaled SSM may need further validation campaigns like field surveys and

768 experiments, based on which the data quality flag can be better built for the product's

769 futural version.

770 The methodological framework proposed in this paper is prospective to be

771 universally applied in other regions of the world to serve for better monitoring of the

772 global surface wetness in the following studies. If applied in continental and global

773 scales, however, the current process for gap-filling of PM seams may require further

774 attention and improvement. In this study, SSM in regions intersected with PM-seam

775 gaps were estimated using TB observations from PM swaths at neighboring dates (see

776 Equation-5). Although the errors in the PM-seam gaps over China as reported by Table

777 2 are only slightly larger compared to the PM-covered regions, they cannot be ignorable

778 completely and may leave extra concern on the universality of this technique, especially

779 in the low latitudinal tropical regions where the effect of PM-seam gap is more apparent

780 than in our study area. Besides, another imperfection of this data product lies in the gap

781 period between AMSR-E and AMSR-2. Considering the different systematic error

782 patterns of various PM SSM products, we did not generate downscaled SSM based on

783 other PM products (e.g. the SMOS SSM product) during this period but just left the

784 period as null values. We suggest a more rigorous and universal inter-calibration

785 framework on different PM SSM products to be developed in the future for a long-term

786 consistent 1-km downscaled SSM dataset.

5. Conclusions

787

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

808 hydrological processes and water resource management.
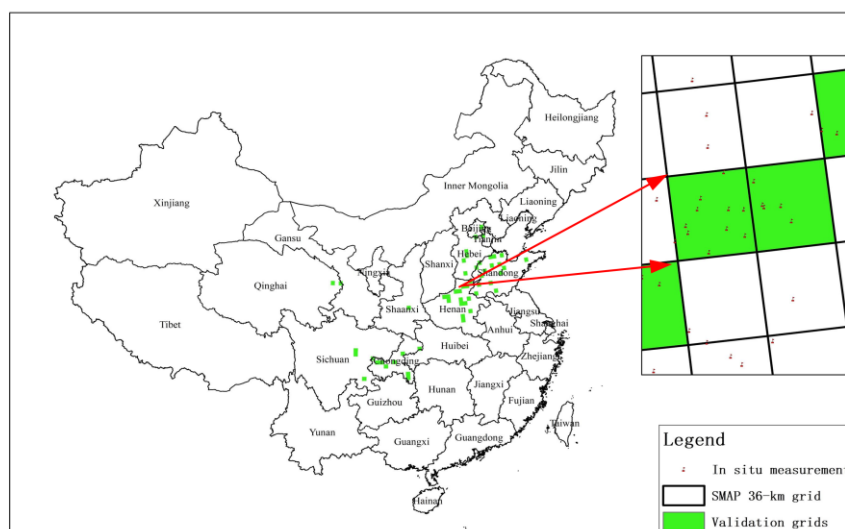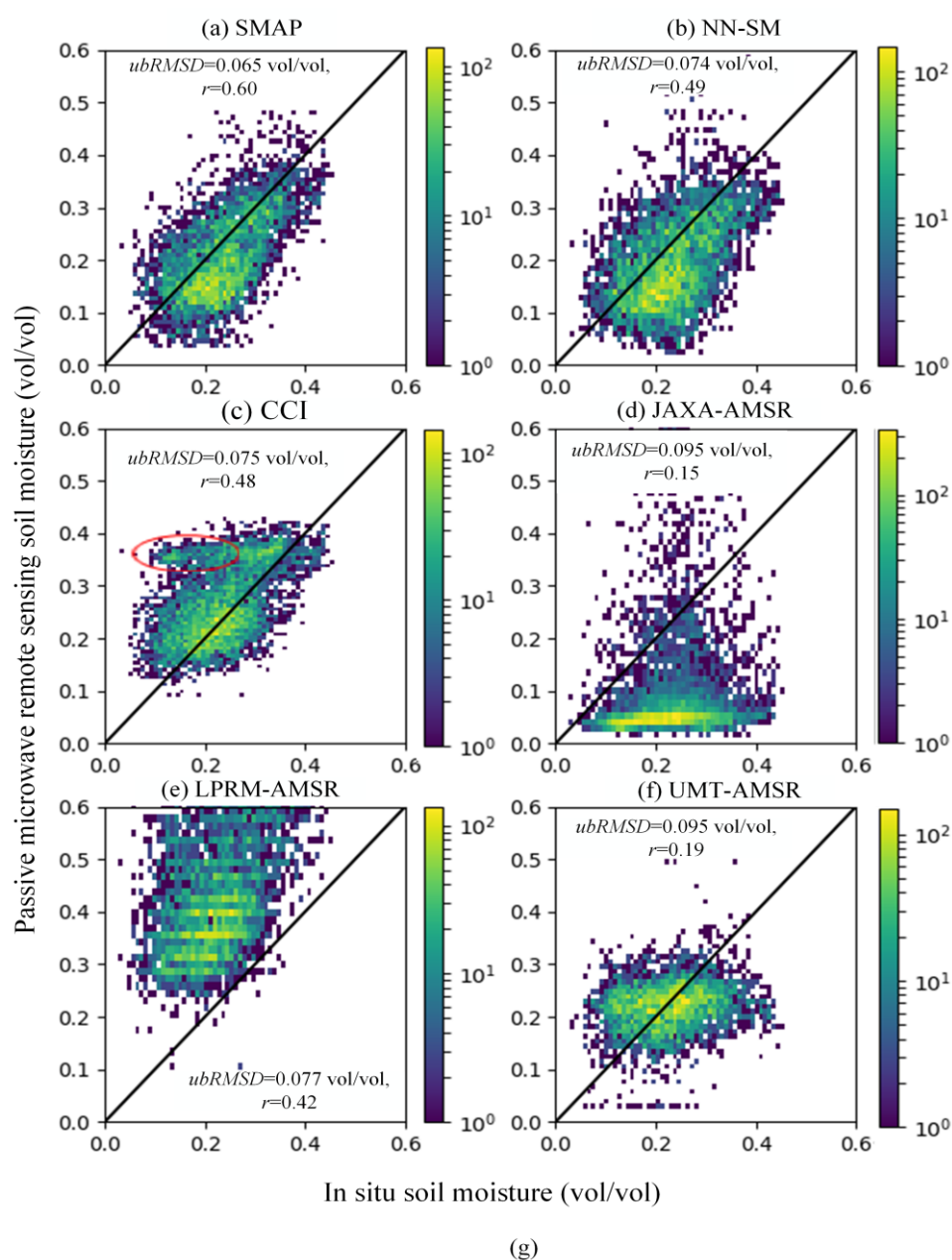
809

Appendix

A. Evaluation on different PM SSM products

We have made evaluations on the various AMSR-based SSM products (as shown in Table 1) covering the recent 10 years or longer, based on our soil moisture observation network all over China. The 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-

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

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

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

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

(a) SMAP
ubRMSD=0.065 vol/vol, *r*=0.60

(b) NN-SM
ubRMSD=0.074 vol/vol, *r*=0.49

(c) CCI
ubRMSD=0.075 vol/vol, *r*=0.48

(d) JAXA-AMSR
ubRMSD=0.095 vol/vol, *r*=0.15

(e) LPRM-AMSR
ubRMSD=0.077 vol/vol, *r*=0.42

(f) UMT-AMSR
ubRMSD=0.095 vol/vol, *r*=0.19

Passive microwave remote sensing soil moisture (vol/vol)

In situ soil moisture (vol/vol)

(g)

Legend
In situ measurement
SMAP 36-km grid
Validation grids

835

Fig. A1 (a)-(f) Comparison of different PM SSM products (as reported in 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, 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

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

859   Fig. A2, the bias adjusted cloudy LST with better gap-filling accuracies, however,

860   obtained inferior performance in SSM downscaling. This final validation result, to

861   some degree, confirms our assumption in Section 2.2.2 that the reconstructed cloudy

862   LST but for the hypothesized clear-sky condition is the better proxy of surface moisture

863   dynamics. But overall, as all LST estimates discussed herein are for the midnight

864   scenario (when the energy interaction between atmosphere and land surface is relatively

865   weak), the RMSD difference for different weather conditions in Fig.A2 is expectedly

866   marginal. As a consequence, the difference in ubRMSD of SSM in Fig.A3 can hardly

867   be identified as 'very significant'. Therefore, we encourage further tests on this

868   conclusion in specific future studies to confirm its universality, especially for situation

869   of the 'morning to noon' time window.

870
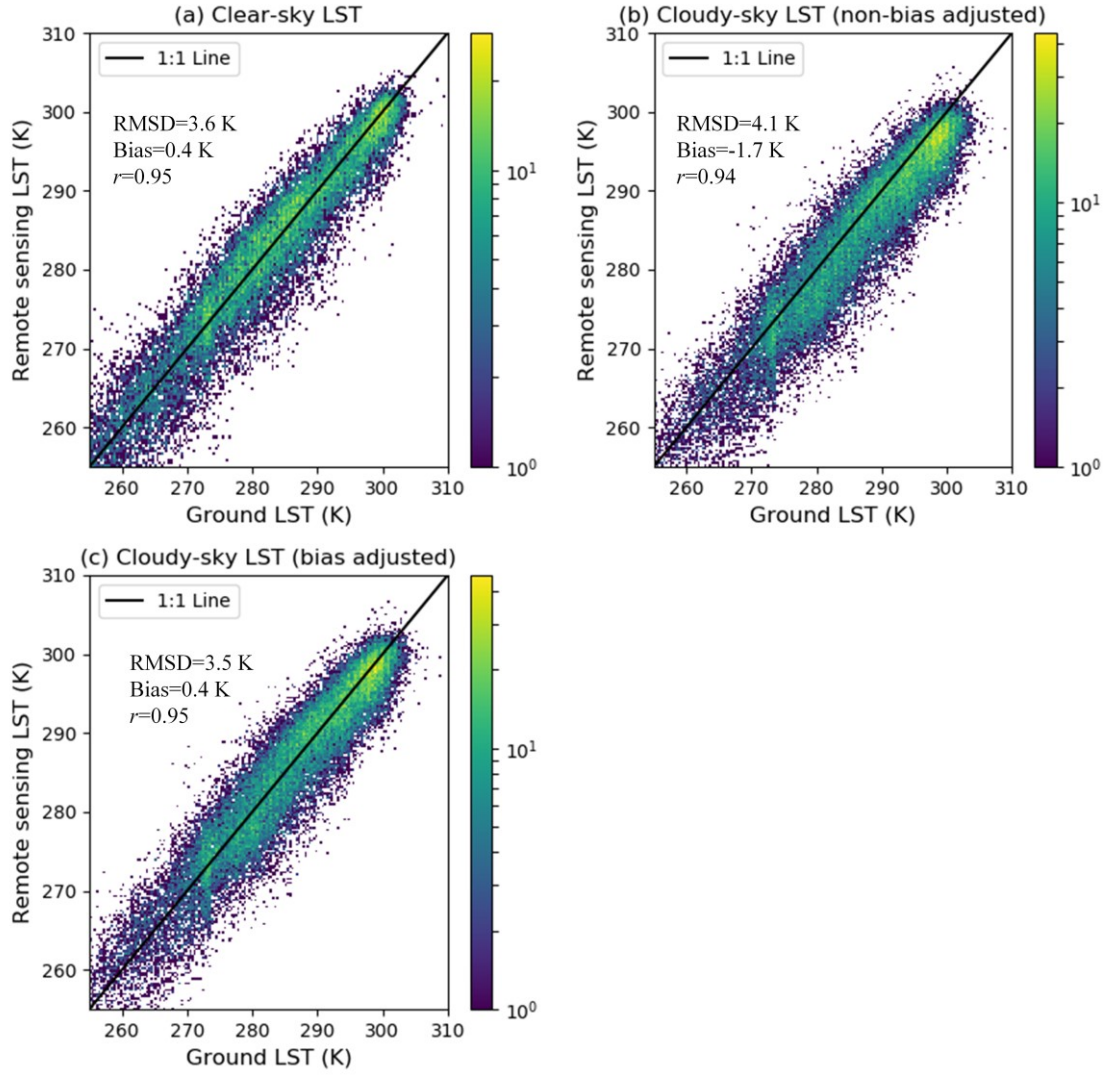
Fig. A2 Validation of the clear sky LST (a), reconstructed LST under cloud but with no passive-

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

microwave based bias adjustment (c) respectively, based on the 0-cm ground temperature
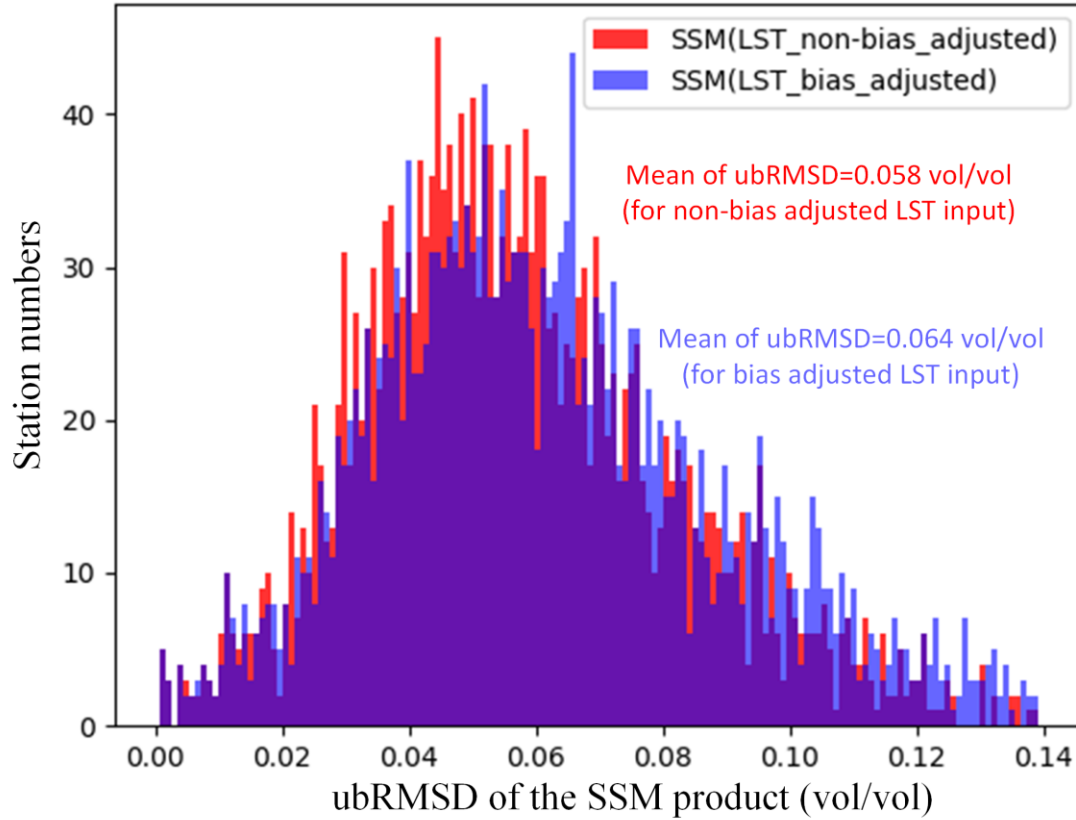
measurements at meteorological stations.

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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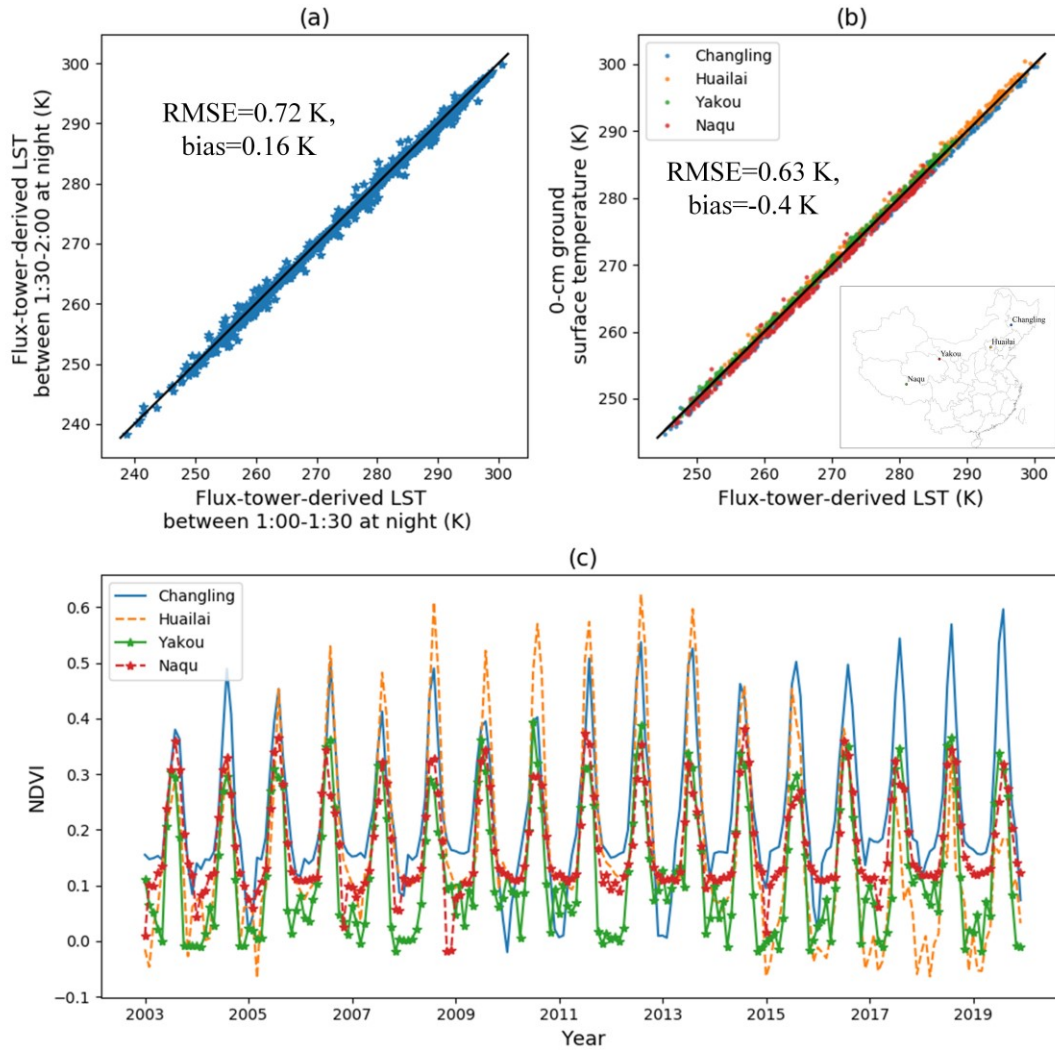
908 thermal infrared LST.

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

56

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

The authors declare that they have no conflict of interest.

Data availability

The published SSM dataset is available under the Creative Commons Attribution

4.0 International License at the following link:

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.

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

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

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

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

971 10.5194/hess-25-5749-2021, 2021.

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

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

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

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

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

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

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

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

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

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

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

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

984 H24C-08,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

1000   2013.

1001   Jeffrey, P., Walker, Paul, R., and Houser: A methodology for initializing soil moisture in a global climate model:

1002   Assimilation of near-surface soil moisture observations, Journal of Geophysical Research Atmospheres, 2001.

1003   Jiménez, C., Prigent, C., Ermida, S. L., and Moncet, J. L.: Inversion of AMSR-E observations for land surface

1004   temperature estimation: 1. Methodology and evaluation with station temperature, Journal of Geophysical Research:

1005   Atmospheres, 2017.

1006   Jing, Z. and Zhang, X.: A soil moisture assimilation scheme using satellite-retrieved skin temperature in meso-scale

1007   weather forecast model, Atmos Res, 95, 333-352, 2010.

1008   Jones, L. A., Kimball, J. S., Podest, E., McDonald, K. C., Chan, S. K., and Njoku, E. G.: A method for deriving land

1009   surface moisture, vegetation optical depth, and open water fraction from AMSR-E, IEEE IGARSS 2009., Cape Town,

1010   South Africa, 2009, III-916-III-919, http://doi.org/10.1109/IGARSS.2009.5417921,

1011   Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., . . . Zhang, K.: Recent decline in

1012   the global land evapotranspiration trend due to limited moisture supply, Nature, 467, 951-954, 10.1038/nature09396, 2010.

1013   Kim, J. and Hogue, T. S.: Improving spatial soil moisture representation through integration of AMSR-E and MODIS

1014   products, IEEE Trans. Geosci. Remote Sens, 50, 446-460, http://doi.org/10.1109/TGRS.2011.2161318, 2012.

1015   Koike, T., Nakamura, Y., Kaihotsu, I., Davva, G., Matsuura, N., Tamagawa, K., and Fujii, H.: Development of an

1016   Advanced Microwave Scanning Radiometer (AMSR-E) algorithm of soil moisture and vegetation water content (written

1017   in Japanese), Annual Journal of Hydraulic Engineering, 48, 217-222 2004.

1018   Komatsu, T. S.: Toward a Robust Phenomenological Expression of Evaporation Efficiency for Unsaturated Soil

1019   Surfaces, Journal of Applied Meteorology, 42, 1330-1334, 10.1175/1520-0450(2003)042<1330:Tarpeo>2.0.Co;2, 2003.

1020   Kong, D., Zhang, Y., Gu, X., and Wang, D.: A robust method for reconstructing global MODIS EVI time series on the

1021   Google Earth Engine, Isprs J Photogramm, 155, 13-24, 2019.

1022   Koster, R. D., Mahanama, S., Livneh, B., Lettenmaier, D. P., and Reichle, R. H.: Skill in streamflow forecasts derived

1023   from large-scale estimates of soil moisture and snow, Nature Geoscience, 3, 613-616, 2010.

1024   Malbéteau, Y., Merlin, O., Molero, B., Rüdiger, C., and Bacon, S.: DisPATCh as a tool to evaluate coarse-scale remotely

1025   sensed soil moisture using localized in situ measurements: Application to SMOS and AMSR-E data in Southeastern

1026   Australia, Int J Appl Earth Obs, 45, 221-234, https://doi.org/10.1016/j.jag.2015.10.002, 2016.

1027   Meesters, A. G. C. A., De Jeu, R. A. M., and Owe, M.: Analytical derivation of the vegetation optical depth from the

1028   microwave polarization difference index, IEEE Geosci. Remote Sens. Lett., 2, 121-123, 2005.

1029   Mendoza, P. A., Mizukami, N., Ikeda, K., Clark, M. P., Gutmann, E. D., Arnold, J. R., . . . Rajagopalan, B.: Effects of

1030   different regional climate model resolution and forcing scales on projected hydrologic changes, J. Hydrol., 541, 1003-1019,

1031   https://doi.org/10.1016/j.jhydrol.2016.08.010, 2016.

1032   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

1033   dataset for China in 2002-2018, Earth System Science Data, 13, 3239-3261, 10.5194/essd-13-3239-2021, 2021.

1034   Merlin, O., Al Bitar, A., Walker, J. P., and Kerr, Y.: An improved algorithm for disaggregating microwave-derived

1035   soil moisture based on red, near-infrared and thermal-infrared data, Remote Sens. Environ., 114, 2305-2316,

1036   https://doi.org/10.1016/j.rse.2010.05.007, 2010.

1037   Merlin, O., Walker, J. P., Chehbouni, A., and Kerr, Y.: Towards deterministic downscaling of SMOS soil moisture

1038   using   MODIS   derived   soil   evaporative   efficiency,   Remote   Sens.   Environ.,   112,   3935-3946,

1039   http://doi.org/10.1016/j.se.2008.06.012, 2008.

1040   Merlin, O., Chehbouni, A. G., Kerr, Y. H., Njoku, E. G., and Entekhabi, D.: A combined modeling and

1041   multipectral/multiresolution remote sensing approach for disaggregation of surface soil moisture: Application to SMOS

1042   configuration, IEEE Trans. Geosci. Remote Sens, 43, 2036-2050, http://doi.org/10.1109/TGRS.2005.853192, 2005.

1043   Merlin, O., Escorihuela, M. J., Mayoral, M. A., Hagolle, O., Al Bitar, A., and Kerr, Y.: Self-calibrated evaporation-

1044   based disaggregation of SMOS soil moisture: An evaluation study at 3 km and 100 m resolution in Catalunya, Spain,

1045   Remote Sens. Environ., 130, 25-38, 10.1016/j.rse.2012.11.008, 2013.

1046   Merlin, O., Malbeteau, Y., Notfi, Y., Bacon, S., Er-Raki, S., Khabba, S., and Jarlan, L.: Performance Metrics for Soil

1047   Moisture Downscaling Methods: Application to DISPATCH Data in Central Morocco, Remote Sens., 7, 3783-3807,

1048   http://doi.org/10.3390/rs70403783, 2015.

1049   Molero, B., Merlin, O., Malbéteau, Y., Al Bitar, A., Cabot, F., Stefan, V., . . . Jackson, T. J.: SMOS disaggregated soil

1050   moisture product at 1km resolution: Processor overview and first validation results, Remote Sens. Environ., 180, 361-376,

1051   http://doi.org/10.1016/j.rse.2016.02.045, 2016.

1052   Montaldo, N., Albertson, J. D., Mancini, M., and Kiely, G.: Robust simulation of root zone soil moisture with

1053   assimilation of surface soil moisture data, Water Resour Res, 37, 2889-2900, 10.1029/2000WR000209, 2001.

1054 O'Neill, P. E., Bindlish, R., Chan, S., Chaubell, J., Colliander, A., Njoku, E., and Jackson, T.: SMAP Algorithm

1055 Theoretical Basis Document: Level 2 & 3 Soil Moisture (Passive) Data Products, Revision G., Jet Propulsion Laboratory,

1056 Pasadena, CA, 2021.

1057 Owe, M., de Jeu, R., and Walker, J.: A methodology for surface soil moisture and vegetation optical depth retrieval

1058 using the microwave polarization difference index, IEEE Trans. Geosci. Remote Sens, 39, 1643-1654, 2001.

1059 Pan, H., Chen, Z., Wit, A. D., and Ren, J.: Joint Assimilation of Leaf Area Index and Soil Moisture from Sentinel-1

1060 and Sentinel-2 Data into the WOFOST Model for Winter Wheat Yield Estimation, Sensors (Basel, Switzerland), 19, 2019.

1061 Peng, J., Loew, A., Zhang, S. Q., Wang, J., and Niesel, J.: Spatial downscaling of satellite soil moisture data using a

1062 vegetation temperature condition index, IEEE Trans. Geosci. Remote Sens, 54, 558-566,

1063 http://doi.org/10.1109/TGRS.2015.2462074, 2016.

1064 Piles, M., Entekhabi, D., and Camps, A.: A Change Detection Algorithm for Retrieving High-Resolution Soil Moisture

1065 From SMAP Radar and Radiometer Observations, IEEE Trans. Geosci. Remote Sens, 47, 4125-4131,

1066 10.1109/TGRS.2009.2022088, 2009.

1067 Sabaghy, S., Walker, J. P., Renzullo, L. J., Akbar, R., Chan, S., Chaubell, J., . . . Yueh, S.: Comprehensive analysis of

1068 alternative downscaled soil moisture products, Remote Sens. Environ., 239, 111586,

1069 https://doi.org/10.1016/j.rse.2019.111586, 2020.

1070 Sanchez-Ruiz, S., Piles, M., Sanchez, N., Martinez-Fernandez, J., Vall-Ilossera, M., and Camps, A.: Combining SMOS

1071 with visible and near/shortwave/thermal infrared satellite data for high resolution soil moisture estimates, J. Hydrol., 516,

1072 273-283, 10.1016/j.jhydrol.2013.12.047, 2014.

1073 Scaini, A., Sanchez, N., Vicente-Serrano, S. M., and Martinez-Fernandez, J.: SMOS-derived soil moisture anomalies

1074 and drought indices: a comparative analysis using in situ measurements, Hydrological Processes, 29, 373-383,

1075 10.1002/hyp.10150, 2015.

1076 Song, P. and Zhang, Y.: An improved non-linear inter-calibration method on different radiometers for enhancing

1077 coverage of daily LST estimates in low latitudes, Remote Sens. Environ., 264, 112626,

1078 https://doi.org/10.1016/j.rse.2021.112626, 2021a.

1079 Song, P. and Zhang, Y.: Daily all weather surface soil moisture data set with 1 km resolution in China (2003-2019),

1080 National Tibetan Plateau Data Center [dataset], 10.11888/Hydro.tpdc.271762, 2021b.

1081 Song, P., Huang, J., and Mansaray, L. R.: An improved surface soil moisture downscaling approach over cloudy areas

1082 based on geographically weighted regression, Agr Forest Meteorol, 275, 146-158, 10.1016/j.agrformet.2019.05.022, 2019a.

1083   Song, P., Zhang, Y., and Tian, J.: Improving Surface Soil Moisture Estimates in Humid Regions by an Enhanced

1084   Remote Sensing Technique, Geophys Res Lett, 48, e2020GL091459, https://doi.org/10.1029/2020GL091459, 2021.

1085   Song, P., Mansaray, L. R., Huang, J., and Huang, W.: Mapping paddy rice agriculture over China using AMSR-E

1086   time series data, Isprs J Photogramm, 144, 469-482, 10.1016/j.isprsjprs.2018.08.015, 2018.

1087   Song, P., Huang, J., Mansaray, L. R., Wen, H., Wu, H., Liu, Z., and Wang, X.: An Improved Soil Moisture Retrieval

1088   Algorithm Based on the Land Parameter Retrieval Model for Water-Land Mixed Pixels Using AMSR-E Data, IEEE Trans.

1089   Geosci. Remote Sens, 1-15, 10.1109/TGRS.2019.2915346, 2019b.

1090   Stefan, V. G., Merlin, O., Escorihuela, M.-J., Molero, B., and Er-Raki, S.: Temporal Calibration of an Evaporation-

1091   Based Spatial Disaggregation Method of SMOS Soil Moisture Data, Remote Sens., 12, 1671, 2020.

1092   Sui, D. Z.: Tobler's First Law of Geography: A Big Idea for a Small World?, Annals of the Association of American

1093   Geographers, 94, 269-277, https://doi.org/10.1111/j.1467-8306.2004.09402003.x, 2004.

1094   Ulaby, F. T. and Wilson, E. A.: Microwave Attenuation Properties of Vegetation Canopies, IEEE Trans. Geosci.

1095   Remote Sens, GE-23, 746-753, 10.1109/TGRS.1985.289393, 1985.

1096   Vergopolan, N., Xiong, S. T., Estes, L., Wanders, N., Chaney, N. W., Wood, E. F., . . . Sheffield, J.: Field-scale soil

1097   moisture bridges the spatial-scale gap between drought monitoring and agricultural yields, Hydrol. Earth Syst. Sci., 25,

1098   1827-1847, 2021.

1099   Verstraeten, W. W., Veroustraete, F., van der Sande, C. J., Grootaers, I., and Feyen, J.: Soil moisture retrieval using

1100   thermal inertia, determined with visible and thermal spaceborne data, validated for European forests, Remote Sens.

1101   Environ., 101, 299-314, 2006.

1102   Wang, K. and Liang, S.: Evaluation of ASTER and MODIS land surface temperature and emissivity products using

1103   long-term surface longwave radiation observations at SURFRAD sites, Remote Sens. Environ., 113, 1556-1565,

1104   https://doi.org/10.1016/j.rse.2009.03.009, 2009.

1105   Wang, L. and Qu, J. J.: NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture

1106   with satellite remote sensing, Geophys Res Lett, 34, L20405, 10.1029/2007GL031021, 2007.

1107   Wei, Z., Meng, Y., Zhang, W., Peng, J., and Meng, L.: Downscaling SMAP soil moisture estimation with gradient

1108   boosting decision tree regression over the Tibetan Plateau, Remote Sens. Environ., 225, 30-44, 2019.

1109   Wu, D., Liang, H., Cao, T., Yang, D., Zhou, W., and Wu, X.: Construction of operation monitoring system of automatic

1110   soil moisture observation network in China, Meteorological Science and Technology, 42, 278-282, 2014

1111   Yang, G., Sun, W. W., Shen, H. F., Meng, X. C., and Li, J. L.: An Integrated Method for Reconstructing Daily MODIS

1112   Land Surface Temperature Data, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 12, 1026-1040, 2019.

1113   Yao, P., Lu, H., Shi, J., Zhao, T., Yang, K., Cosh, M. H., . . . Entekhabi, D.: A long term global daily soil moisture

1114   dataset derived from AMSR-E and AMSR2 (2002–2019), Scientific Data, 8, 143, 10.1038/s41597-021-00925-8, 2021.

1115   Zeng, Y., Feng, Z., and Xiang, N.: Assessment of soil moisture using Landsat ETM+ temperature/vegetation index in

1116   semiarid environment, IEEE International Geoscience & Remote Sensing Symposium, Piscataway NJ, 2004, 4306-4309

1117   vol.4306, 10.1109/IGARSS.2004.1370089,

1118   Zhang, J., Zhou, Z., Yao, F., Yang, L., and Hao, C.: Validating the Modified Perpendicular Drought Index in the North

1119   China Region Using In Situ Soil Moisture Measurement, IEEE Geoscience & Remote Sensing Letters, 12, 542-546, 2014.

1120   Zhang, Y., Kong, D., Gan, R., Chiew, F. H. S., Mcvicar, T. R., Zhang, Q., and Yang, Y.: Coupled estimation of 500 m

1121   and 8-day resolution global evapotranspiration and gross primary production in 2002-2017, Remote Sens. Environ., 222,

1122   165-182, 2019.

1123   Zhang, Y. Q., Chiew, F. H. S., Liu, C. M., Tang, Q. H., Xia, J., Tian, J., . . . Li, C. C.: Can Remotely Sensed Actual

1124   Evapotranspiration Facilitate Hydrological Prediction in Ungauged Regions Without Runoff Calibration?, Water Resour

1125   Res, 56, 2020.

1126   Zheng, J. Y., Lu, H. S., Crow, W. T., Zhao, T. J., Merlin, O., Rodriguez-Fernandez, N., . . . Gou, Q. Q.: Soil moisture

1127   downscaling using multiple modes of the DISPATCH algorithm in a semi-humid/humid region, Int J Appl Earth Obs, 104,

1128   10.1016/j.jag.2021.102530, 2021.

1129   Zhou, S., Williams, A. P., Lintner, B., Berg, A. M., and Gentine, P.: Soil moisture–atmosphere feedbacks mitigate

1130   declining water availability in drylands, Nature Climate Change, 11, 2021.

1131   Zhu, Z. and Shi, C.: Simulation and Evaluation of CLDAS and GLDAS Soil Moisture Data in China (written in

1132   Chinese), Science Technology and Engineering, 32, 138-144, 2014.

1133