Estimate of Cloudy-Sky Surface Emissivity From Passive Microwave Satellite Data Using Machine Learning

Xin-Ming Zhu, Xiao-Ning Song, Pei Leng, Zhao-Liang Li, Xiao-Tao Li, Liang Gao, Da Guo

Estimate of Cloudy-sky Surface Emissivity from Passive Microwave Satellite Data Using Machine Learning

Xin-Ming Zhu, Xiao-Ning Song, Pei Leng, Zhao-Liang Li, Xiao-Tao Li, Liang Gao, and Da Guo

***Abstract*—The derivation of microwave land surface emissivity (MLSE) under various weather conditions from the microwave radiometer plays a crucial role in acquiring land surface and atmospheric parameters. Nevertheless, currently, most existing studies mainly focus on the clear-sky scenarios owing to a lack of cloudy-sky land surface temperature (LST) and uncertainties in simulating the scattering and emission properties of atmospheric hydrometeors. Under this background, with satellite observations and the random forest (RF) model, this study proposes a method to estimate the MLSE under cloudy skies. First, clear-sky MLSEs with satisfactory accuracy are retrieved by using the brightness temperatures (BTs) from the Advanced Microwave Scanning Radiometer-Earth sensor, LSTs from the Moderate Resolution Imaging Spectroradiometer, and atmospheric profiles from the ERA5 reanalysis. Then, the relation among the clear-sky MLSE and related impact factors is built with the RF and extended to the cloudy-sky environment for generating all-weather MLSEs with a 0.25°. The results show that the input datasets present a considerable impact on the calculation of instantaneous MLSE, and a 5.73 K bias of ERA5 LST may generate a 0.014-0.021 error in the MLSE from 6.9 to 89 GHz horizontal polarization, while the impacts of BT and profile uncertainties on the MLSE are smaller. The retrieved clear-sky MLSE is coincident with the existing MLSE for the spatiotemporal variations, and there is an average difference range from -0.035 to 0.035 in January 2008. Meanwhile, the constructed RF model can successfully apply to cloudy-sky status and recover the MLSE image gaps affected by cloud contamination.**

***Index Terms*—All-weather, AMSR-E product, error analysis, microwave land surface emissivity, random forest.**

Manuscript received ; revised ; accepted . This work was supported in part by the National Natural Science Foundation of China under Grant 41871242 and 42041005, the Fundamental Research Funds for the Central Universities, and the China scholarship council. *(Corresponding author: Xiao-Ning Song)*

X.-M. Zhu is with College of Resources and Environment, University of the Chinese Academy of Sciences, Beijing 100049, China; ICube (UMR 7357), UdS, CNRS, 300 Bld Sebastien Brant, CS 10413, Illkirch 67412, France (e-mail: zhuxinming19@mails.ucas.ac.cn).

X.-N. Song is with College of Resources and Environment, University of the Chinese Academy of Sciences, Beijing 100049, China; Yanshan Earth Critical Zone and Surface Fluxes Research Station, University of Chinese Academy of Sciences, Beijing 101408, China (e-mail: songxn@ucas.ac.cn).

P. Leng is with the Key Laboratory of Agricultural Remote Sensing, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China (e-mail: lengpei@caas.cn).

Z.-L. Li is with the ICube (UMR 7357), UdS, CNRS, 300 Bld Sebastien Brant, CS 10413, Illkirch 67412, France; the Key Laboratory of Agricultural Remote Sensing, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China (e-mail: lizl@unistra.fr).

X.-T. Li is with the Chinese Institute of Water Resource and Hydropower Research, Beijing 100038, China (e-mail: lixt@iwhr.com).

L. Gao is with College of Resources and Environment, University of the Chinese Academy of Sciences, Beijing 100049, China (e-mail: gaoliang17@mails.ucas.ac.cn).

D. Guo is with College of Resources and Environment, University of the Chinese Academy of Sciences, Beijing 100049, China (e-mail: guoda18@mails.ucas.ac.cn).

I. Introduction

THE microwave land surface emissivity (MLSE) is a key parameter many researchers use to investigate the physical processes of earth-atmosphere interactions at the regional and global scales [1-3]. It is defined as the ratio of the emitted radiance of the surface material to that of a blackbody at the same temperature level and changes with the frequencies due to the difference in penetration depth of the microwave [1]. In reality, the MLSE variable can be widely applied in many research fields, such as the retrieval of atmospheric variables such as water vapor content (WV), cloud liquid water (CLW), and precipitating water [4-8]; the estimation of soil moisture (SM) [9-11]; the research of vegetation phenology change [12-14]; and the application of multi-source data assimilation in the numerical weather prediction systems [15-18]. As a result, a reliable and large range of estimation of MLSE images under various weather status is of great significance and is conducive to understanding and revealing the water and energy exchange between the land and the atmosphere.

In contrast with the estimation of microwave emissivity on the ocean, the derivation of MLSE is more difficult due to its high temporal variability and spatial heterogeneity caused by the complex surface geophysical characteristics and geometric morphology, which include the SM, soil texture, land surface roughness, land-cover type, and vegetation optical depth, etc. [19, 20]. For a better understanding of surface parameters and atmospheric properties over land regions, over the past three decades, considerable progress has been made in estimating the MLSE using three main methods: the physical modeling calculation, one-dimensional variational retrieval, and satellite observation-based retrieval [2, 3, 18, 19, 21-31]. Among them, the physical model-based estimation primarily simulates the MLSE data by modeling the interaction between land surface components and the electromagnetic wave; but, these models require abundant auxiliary parameters that are not readily available or observable at continental or global scales, e.g., the surface roughness, soil dielectric properties, forest structures,

and snow grain sizes, etc. [20]. Representative models mainly include the Advanced Integral Equation Model (AIEM) proposed by Chen et al. [22], the LandEM embedded in the Community Radiative Transfer Model (CRTM) proposed by Weng et al. [19], and the Community Microwave Emission Model (CMEM) developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) [16, 29]. The one-dimensional variational retrieval method mainly uses the Bayes principle to obtain an optimal estimation of MLSEs by blending the observed background fields from various data sources, whereas it is necessary to establish an accurate initial guess of surface emissivity. For instance, the ECMWF used the Special Sensor Microwave Imager (SSM/I) MLSE dataset as the initial guess to propose a calculation tool, i.e., A Tool to Estimate Land Surface Emissivities at Microwave Frequencies (TELSEM) [18, 28]. Compared with the above two methods, the satellite observation-based retrieval takes advantage of the passive microwave (PMW) satellite data to obtain the spatially continuous MLSEs over a large range. In general, this method is based on the radiative transfer equation (RTE) and makes full use of BTs at the top of the atmosphere (TOA), land surface temperatures (LSTs), and atmospheric profiles. Thus, this approach possesses a reliable physical meaning and has become the mainstream for obtaining the MLSE nowadays.

More recently, many scholars have employed the satellite retrieval method to conduct research on the inversion of microwave emissivity and have developed a series of MLSE products [2, 6, 26, 30-34]. For instance, by assuming that the atmospheric conditions under a clear sky status are transparent, Wilke and Mcfarland [35] estimated the MLSE with the ratio of BTs and LSTs. Nevertheless, in consideration of the evident contributions of atmospheric properties to the BTs at the TOA, some more accurate retrieval approaches were proposed by considering the atmospheric attenuation, radiation, and land surface emissions. Prigent et al. [2] and Karbou et al. [26] yielded MLSE products under clear skies using the BTs from the SSM/I radiometer and the Advanced Microwave Sounding Unit sensor, respectively. In addition, some other radiometers onboard low-Earth orbiting satellites were also applied to estimate MLSEs, such as the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) [33, 36], the AMSR2 [31], the Tropical Rainfall Measuring Mission Microwave Imager [6], the WindSat onboard the Coriolis satellite [37], the Microwave Analysis and Detection of Rain and Atmospheric Structures [38], and the Chinese Microwave Radiation Imager onboard the FY-3B satellite [39]. With these MLSE products, some relevant studies, like the error source analysis of MLSE estimate [40, 41], the comparison among various MLSE datasets [42, 43], and the satellite retrievals of atmospheric and vegetation phenology parameters, have been widely conducted [14, 44]. Despite this progress, the surface emissivity retrieval based on the satellite observation is mostly designed to operate under clear-sky conditions. Few studies used this kind of approach to generate cloudy-sky MLSEs. Generally, there are two key reasons to explain this issue. First, it is relatively difficult to correct the atmospheric contributions from cloudy-sky status due to the lack of cloud/rain properties (e.g., the vertical profiles of cloud water, cloud ice, total rain, total snow, effective cloud fraction) and the uncertainty of the atmospheric radiation transmission simulation, which directly controls the atmospheric correction in all-sky situation [30]. In particular, for higher frequencies, the atmospheric absorption, emission, and scattering of surface signals are stronger, and a large amount of uncertainty will arise in this case. Second, currently, the estimates of MLSEs mostly depend on clear-sky LSTs derived from the thermal infrared (TIR) bands, like the Moderate Resolution Imaging Spectroradiometer (MODIS) LST because there are no adequate all-weather LST dataset. Although some reanalysis datasets provided all-weather skin temperatures, the accuracies of these data and their impacts on MLSEs still remain problematic, especially at cloudy skies.

Recently, several researchers have estimated the cloudy-sky MLSEs. By assuming that the difference between the MLSEs under clear and cloudy skies can be ignored, Aires et al. [23] used a neural network approach to retrieve surface emissivities between 19 and 85 GHz from the SSM/I BTs and clear-sky MLSE previously calculated. Using the microwave radiative transfer model, Lin and Minnis [45] combined the SSM/I BTs and the ground observation radiances at the southern Great Plains to retrieve the MLSE in both clear and cloudy skies. In addition, Baordo and Geer [46] retrieved all-sky MLSEs using the BTs from the SSM/I and the effective cloud fraction properties from the ECMWF by considering the impacts of clear and cloudy skies on the BTs. Hu et al. [34] estimated the MLSEs of China under both clear and cloudy skies using the AMSR-E BTs, MODIS cloud-layer properties, and reanalysis skin temperatures as well as atmospheric profiles from the ERA-20C with the forward simulation. In spite of this, these works still need to be further investigated. On the one hand, the uncertainties of input datasets (i.e., LST and profiles of temperature and humidity) play a crucial role in the MLSE estimates. On the other hand, because the effects of clouds on the microwave radiation are relatively complicated, using the radiative transfer model to correct the atmospheric effect in the presence of clouds/rain is difficult due to its limitations [47, 48]. Thus, it remains necessary to develop a new scheme to obtain the MLSEs under various weather circumstances.

According to the above statement, the main objective of this study is to construct a method by integrating the advantages of remote sensing retrieval and a machine learning algorithm for acquiring MLSE datasets under clear and cloudy skies. Then, applying this method to the China region in 2008 to reveal the spatiotemporal characteristics of MLSE. Likewise, given that the difference of MLSEs between clear and cloudy skies is not significant [23] and the MLSE sensitivity to the atmospheric properties is less [34], we assume that the relations between the clear-sky MLSEs and related variables could be extended to cloudy-sky conditions. Then, we use the random forest (RF) model to construct the nonlinear relations between the MLSE and related impact factors under clear sky status to produce cloudy-sky MLSEs. We believe that the proposed method can help us to estimate MLSE better and contribute to providing all-weather MLSEs, allowing us to understand the physical processes of Earth-atmosphere interactions.

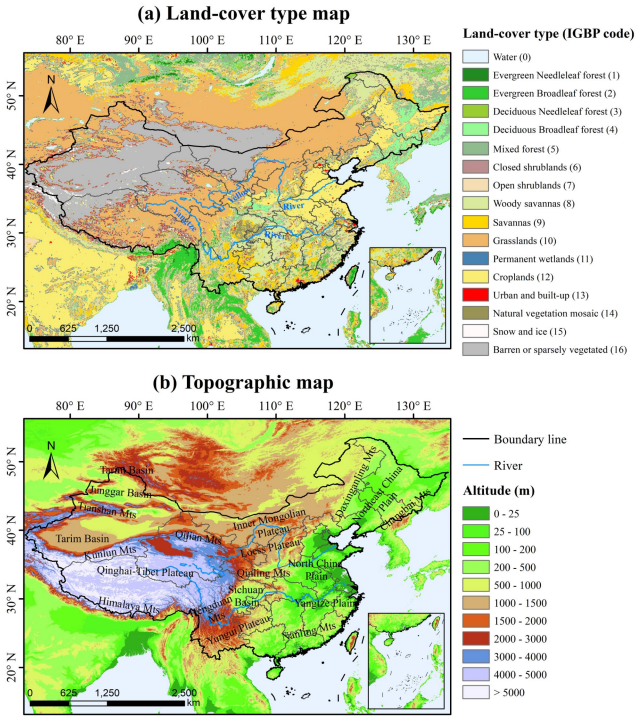


Fig. 1. The land-cover types and terrain of the study area. The land-cover types are derived from the MODIS land-cover product (the classification standard belongs to the international geosphere-biosphere programme, IGBP) and altitude is acquired from the Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM).

The structure of this paper is organized as follows: Section II describes the study area and the used datasets. Section III summarizes the MLSE estimation method in clear and cloudy skies. Section IV presents the impacts of input data sources on clear-sky MLSE retrieval and displays the performances of clear and cloudy skies MLSE estimates. Section V discusses the possible factors impacting MLSE change, the sensitivity of cloud characteristics to the MLSE, the difference between the TIR-derived emissivity and MLSE, and the limitations of the proposed method. Section VI provides the overall conclusions.

II. Study Area and Dataset Used

A. Study Area

A wide range of cloud and rain coverage in China hinders the acquisition of cloudy-sky MLSEs in this region, especially in southern China. Thus, this paper selected China as the study area to implement the estimation of MLSEs under clear and cloudy skies. China is situated in the eastern part of Eurasia and the western part of the Pacific Ocean. The study region includes diverse land-cover types. The topography mainly contains plateaus, mountains, hills, basins, and plains, and there is a pronounced terrain change trend, i.e., high in the west and low in the east. Obviously, the unique geographical location and topographical features contribute to the diversity of China's climate, resulting in a particular distribution of high precipitation in the southeast and a low precipitation rate in the northwest. Fig. 1 presents the land-cover types and terrain of the study area.

TABLE I
The Data Collections Used in This Study

	Data collection	Parameter	Spatial resolution	Temporal resolution
Dataset used for clear-sky MLSE calculation	AMSR-E/Aqua L2A	Brightness temperature (BT)	0.25°	Daily
	MODIS/Aqua MYD11_A1	Land skin temperature (LST)	1-km	Daily
	ERA5 hourly data on pressure levels	Geopotential height, pressure, air temperature, relative humidity, skin temperature	0.25°	Hourly
Dataset used for cloudy-sky MLSE estimation	ERA5 hourly data on single levels	Leaf area index (LAI), soil moisture (SM)	0.25°	Hourly
	SRTM DEM	Elevation, slope, surface roughness	30-m	Statical
	Soil texture	Sand, silt, clay	1-km	Statical
	AMSR-E/Aqua L2A	PR, MPDI	0.25°	Daily
Dataset used for error analysis of MLSE	UWYO sounding profile	Geopotential height, pressure, air temperature, relative humidity	—	Daily
	ERA-Interim reanalysis	Geopotential height, pressure, air temperature, relative humidity, skin temperature	0.75°	6 hourly
	JRA-55 reanalysis		1.25°	6 hourly
	MERRA2 reanalysis		0.25°	3 hourly
	NCEP/FNL reanalysis		1.0°	6 hourly
	NCEP/GFS reanalysis		0.5°	6 hourly
Dataset used for MLSE comparison	Qiu_MLSE	Clear-sky MLSE	0.25°	Daily
	MODIS/Aqua MYD11_A1	Clear-sky TIR-derived emissivity	1-km	Daily

B. Data Collection Used

The data collections used in this study mainly contain: 1) the input datasets used for clear-sky MLSE retrieval, 2) the input datasets used for cloudy-sky MLSE calculation, 3) the datasets used for error analysis of MLSE, and 4) the data used for comparison of MLSE results. The overall characteristics of the datasets used in this study are shown in Table I.

1) Input datasets used for clear-sky MLSE calculation

In order to retrieve clear-sky MLSE with the microwave RTE, it is necessary to acquire the microwave BTs at the TOA, LSTs, and atmospheric profiles of air temperature, humidity, geopotential height, etc. As the AMSR-E and MODIS sensors are onboard the Aqua satellite with good temporal consistency between the obtained datasets in revealing the atmospheric and terrestrial conditions, this study used the AMSR-E/Aqua L2A Global Swath Spatially Resampled BT product and MODIS/Aqua MYD11A1 LST product as main inputs. The Aqua is a scientific research satellite in orbit around the Earth that studies the atmospheric precipitation, surface evaporation, and cycling of water. It probes the Earth twice per day, and the transit times are at approximately 1:30 a.m. (descending orbit) and 1:30 p.m. (ascending orbit) local time at the equator.

The AMSR-E L2A BT product contains six channels at 6.9, 10.6, 18.7, 23.8, 36.5, and 89.0 GHz, with the horizontal (H) polarization and vertical (V) polarization at each channel. Its instantaneous field of view varies from 5.4 km at 89.0 GHz to 56 km at 6.9 GHz, and the Earth's incidence angle is 55° [49]. To match the resolutions of the higher frequency observations to those of the lower frequencies, we employed the resampled datasets having the closest location to the original satellite detection footprint and reprojected them to a 0.25° grid that is equidistant at the equator [50]. The AMSR-E/Aqua L2A data can be freely obtained from the National Snow and Ice Data Center (NSIDC) (https://nsidc.org/data).

The MYD11A1 LST is derived from channels 31 and 32 of MODIS imagery using the generalized split-window algorithm and is of the spatial resolution of 1 km [51]. A previous study has demonstrated that the estimated LSTs have good accuracy of less than 1 K for most homogeneous surfaces and have been widely applied in LST-associated studies [52]. In this study, MYD11A1 LSTs are downloaded from the NASA EARTH DATA website (https://search.earthdata.nasa.gov/).

Meanwhile, we used the clear-sky profiles from the ERA5 reanalysis as input of the atmospheric radiation transfer model (i.e., Monochromatic Radiative Transfer Model, MonoRTM 5.4) for calculating the transmittance (τ), atmospheric upward radiance (T_u), and downward radiance (T_d) values [48]. The ERA5 data has good accuracy in revealing actual atmospheric status [53], and it fuses vast amounts of historical observation records into global estimates with advanced modeling and data assimilation systems. This dataset is generally organized with a regular latitude-longitude grid at a 0.25°×0.25° resolution in GRIB or NetCDF format and has 37 pressure levels within an hourly temporal resolution [54]. Here, the ERA5 profile fields used at clear-sky conditions include the geopotential height, pressure, air temperature, and relative humidity (RH). These data can be derived from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/cdsapp#!/home).

2) Input datasets used for cloudy-sky MLSE estimate

Since the spatiotemporal change of MLSE greatly depends on the characteristics of vegetation, soil water content, soil structure, and terrain factors, the datasets used for cloudy-sky MLSE estimate in this paper mainly used the leaf area index (LAI), SM, soil texture, surface roughness, elevation, slope, and BT indexes that can be applied in the RF training and input. In this study, LAI and SM datasets are provided by the ERA5-Land hourly data, with a 0.25° grid resolution. The LAI refers to one-half of the total green leaf area per unit horizontal ground surface area for the high vegetation type, and the SM belongs to the volume of water in soil layer 1 (0-7 cm) of the ECMWF Integrated Forecasting System. The soil texture data is acquired from the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (http://www.resdc.cn) and has a spatial resolution of 1 km. The soil texture is defined by the composition of particle size, namely, sand, silt, and clay. Elevation, slope, and roughness data are calculated by using the digital elevation model (DEM) derived from the Shuttle Radar Topography Mission (SRTM) dataset, with a 30-m resolution [55].

In addition to the abovementioned datasets, since the heavy precipitation, snow particle, and water body present evident scattering effects on the estimation accuracy of MLSE [56], we also adopted the total column rainwater, snow cover, and land-sea mask datasets derived from the ERA5 hourly data on single levels to filter these regions. They all have a 0.25° grid resolution.

3) Datasets used for error analysis of MLSE estimate

In order to perform the error analysis of MLSE estimation, the corresponding reanalysis profiles and skin temperatures derived from five types of datasets, including the ERA-Interim, the Japanese 55-year Reanalysis Data (JRA-55), the Modern-Era Retrospective analysis for Research and Application versions2 (MERRA2), the National Centers for Environmental Prediction (NCEP)/Final Operational Global Analysis (FNL), and the NCEP/Global Forecasting System (GFS), are also employed to calculate AMSR-E MLSEs under clear skies. After, the clear-sky MLSEs retrieved with the true air-sounding profiles derived from the University of Wyoming (hereafter UWYO profile), AMSR-E BTs, and MODIS LSTs are used as the reference values to reveal the impacts of different input data sources on MLSE calculation.

The introduction of each type of reanalysis data is presented in detail in Table I. The UWYO profile is acquired at 0 and 12 Universal Time Coordinated (UTC) time, and the atmospheric fields include geopotential heights, pressures, air temperatures, and RH. As the Aqua overpass is at local times of 1:30 a.m. and 1:30 p.m every day, all datasets in this section were extracted on UTC 12 at 23 sounding stations (with longitudes ranging from 15°E to 30°E) to guarantee the spatiotemporal correspondence between UWYO profiles and AMSR-E BTs. Detailed station information can refer to [53]. However, it is worth mentioning that all the reanalysis datasets have various atmospheric fields. For comparison purposes, we adopted the same field information as the UWYO profile and processed them in the same way as the ERA5 profile.

4) Datasets used for comparison of MLSE result

Due to the lack of actual *in-situ* MLSE observations, the clear-sky MLSEs estimated in this study are evaluated using the existing MLSE product. With the microwave BTs from the AMSR-E/Aqua L2A, versions 2, the MYD11_L2 LSTs from the MODIS/Aqua, and the MYD07_L2 atmospheric profiles (20 pressure levels) from the MODIS/Aqua, Qiu et al. [36] calculated the daily instantaneous MLSEs from 2002 to 2011 under clear-sky conditions over the global scale (hereafter Qiu_MLSE). This dataset has been compared and tested with the MLSE product of Karbou et al. [26] and can reveal actual microwave radiance conditions from the surface. In this paper, Qiu_MLSE can be freely acquired from the National Tibetan Plateau Data Center of the Institute of Tibetan Plateau Research, Chinese Academy of Sciences.

In addition, for further comparing the difference between the TIR emissivity and MLSE in terms of spatial distribution, we also used the MYD11A1 emissivity product in the MODIS

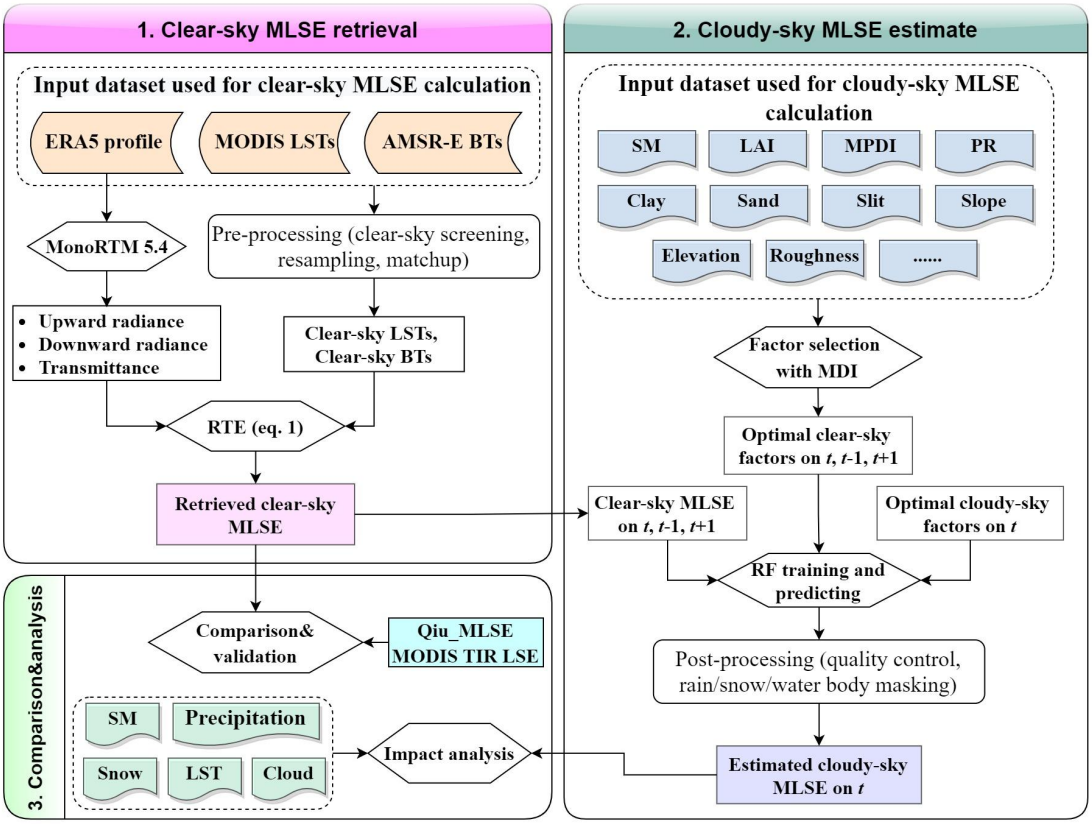


Fig. 2. The overall flowchart of the proposed method.

bands 31 and 32. These surface emissivity datasets are mainly produced with the classification-based look-up table method, which is derived from the land-cover type and its dynamic and seasonal factors [57].

III. Methodology

The implementation flowchart of the proposed method is depicted in Fig. 2. The overall structure can be divided into three sections: the retrieval of clear-sky MLSE, the estimation of cloudy-sky MLSE, and the result comparison.

A. Retrieval of Clear-sky MLSE

In general, the satellite observation-based MLSE retrieval is mainly performed using the microwave RTE. Regarding the non-scattering, plane-parallel atmosphere circumstances, the BTs received by the radiometer for nonblack body boundaries can be expressed as three processes from the PMW sensor to the land surface: (i) radiation information emitted from the land surface and attenuated by the atmosphere; (ii) upward radiation emitted by the atmosphere; and (iii) downward radiation from atmosphere component and cosmic background reflected by the land surface and attenuated by the atmosphere [58]. Based on the Rayleigh-Jean approximation, the MLSE at a specific frequency range can be calculated from satellite observations through the RTE [3]:

$$MLSE(f,\theta)=\frac{T_b(f,\theta)-T_u(f,\theta)-\Gamma(f,\theta)\cdot[T_d(f,\theta)+T_{sky}\cdot\Gamma(f,\theta)]}{T_s\cdot\Gamma(f,\theta)-\Gamma(f,\theta)\cdot[T_d(f,\theta)+T_{sky}\cdot\Gamma(f,\theta)]} \quad (1)$$

where $MLSE(f, \theta)$ is the surface emissivity at frequency f and zenith angle θ; $T_b(f, \theta)$ is the BTs observed by the PMW radiometer at frequency f and zenith angle θ (unit is K); and T_s is the skin temperature, i.e., the effective radiating temperature of the surface at the relevant frequency (unit is K); T_{sky} is the cosmic background radiation (2.75 K).

$\Gamma(f, \theta)$ is the atmospheric optical thickness at frequency f and zenith angle θ, and it is a mathematical function of the atmospheric transmittance τ:

$$\Gamma(f,\theta)=e^{-\tau(f,\theta)} \quad (2)$$

$T_u(f, \theta)$ and $T_d(f, \theta)$ represent the atmospheric upwelling and downwelling radiations emitted by the atmosphere (unit is K), respectively. For an isothermal atmosphere, T_u and T_d can be regarded as approximately equal and are calculated as follows:

$$T_u(f,\theta)=\sec\theta\int_0^H T(z)\kappa(f,z)\exp(-\Gamma(f,z,H)\cdot\sec\theta)dz \quad (3)$$

$$T_d(f,\theta)=\sec\theta\int_0^\infty T(z)\kappa(f,z)\exp(-\Gamma(f,0,z)\cdot\sec\theta)dz \quad (4)$$

where H is the atmospheric height (unit is km); $T(z)$ is the average temperatures of atmospheric components (unit is K); and $\kappa(f, z)$ is the atmospheric attenuation coefficient at the frequency f and height z.

From Eq. (1), we can see the estimation of all-sky MLSE requires an accurate characterization of air temperature, RH, cloud property, and CLW values under all-weather conditions. Another crucial input parameter is all-weather LST. However, currently, it is relatively difficult to acquire all-weather LST and cloud/rain profiles accurately. Thus, the direct retrieval of

MLSE under cloudy skies status is still troublesome. Here, we derived the clear-sky MLSE as the sample dataset of the RF model with Eq. (1) for acquiring cloudy-sky MLSE in the subsequent step. The BTs, LSTs, and atmospheric profiles are obtained from the AMSR-E BTs, MODIS LSTs, and ERA5 reanalysis data, respectively.

Before the MLSE under clear-sky conditions is retrieved, the pre-processing has been implemented. All image pixels are determined to be cloud-free status using collocated MODIS 1-km resolution cloud mask information and the QC file of MYD11A1. Pixels are designated clear skies only when the MODIS algorithm determines "Confident Clear" over the full extent of the largest AMSR-E footprint size. Additionally, to match the resolution between AMSR-E and MODIS datasets, the MODIS LSTs were spatially aggregated to a 0.25° grid by using an average aggregation algorithm. Since τ, T_u, and T_d are not directly available, the MonoRTM 5.4 model developed by Atmospheric & Environmental Research (AER) is used to calculate these three parameters [48] after the transit time and land surface elevation of the ERA5 profile were interpolated using the data processing method proposed by Zhu et al. [53].

B. Prediction of Cloudy-sky MLSEs

Although it is currently knotty to estimate the cloudy-sky MLSE using the satellite observation-based MLSE retrieval method, we can still implement this work by assuming the relation between clear-sky MLSE and related impact variables can be extended to cloudy-sky regions. The main cause is the MLSE, as a unique parameter of the land surface, is primarily influenced by surface properties and rarely by the atmosphere status (in addition to the precipitation) [34, 59, 60]. In addition, there is no significant difference in the MLSE between clear and cloudy skies, according to the research of Aires et al. [23]. Since the RF model has relatively high performance and can automatically settle the nonlinear relationships between the independent variable and dependent variables in the data sets compared with the traditional regressions [61, 62], it is applied here to perform the training and result predictions of MLSE. The basic component of RF regression is the decision tree, and it mainly drives the data by constructing and averaging a large number of randomized and decorrelated trees [61]. This algorithm is insensitive to multiple collinearities of data sets and can effectively avoid the overfitting of a regression model by comprising multiple independent decision trees in the learning procedure [61]. Finally, the RF-based predictions are calculated for each regression tree separately, and then an arithmetic average of the trees as the final forecast result is performed.

The basic equation used for describing the RF prediction for regression results based on the constructed trees is presented as follows [61]:

$$F(x)=\frac{\sum_{j=1}^{N} T_j(x)}{N} \tag{5}$$

where N represents the number of decision trees; T_j denotes each tree; F is a prediction at a new point x as an averaged prediction based on the constructed decision trees.

Before the RF estimation model is created, the selection of the independent factor combination is crucial for the physical parameter prediction [63]. We initially selected the following variables as auxiliary factors to perform the training of RF model due to their interactions with the MLSE: LAI, SM, sand, silt, clay, surface elevation, slope, and land surface roughness. Meanwhile, two AMSR-E BT indices, i.e., the microwave polarization difference index (MPDI, $(BT_V\text{-}BT_H)/(BT_V\text{-}BT_H)$) and the polarization ratio (PR, BT_H/BT_V) at the corresponding frequencies, were used as two key predictors because of their strong correlation with the MLSE and insensitivity to cloudy skies [64, 65]. However, it is worth noting that we did not consider the land-cover types in the RF model training work, even if it also impacts the change of MLSE. This is since the above-selected impact variables more and less contain some information of land-cover types. To prevent the model from being too complex, thus leading to a overfitting issue during the model construction, we adopted the variable importance score of each variable to select the optimal MLSE predictors. Here, the mean decrease in impurity (MDI) is used as the variable importance measure index because it possesses high efficiency and reliability [66]. Via the incessant training, the variables with a low MDI value were removed until the RF model residual reached the minimum. Fig. 3 presents MDI% of ten critical variables at six AMSR-E frequencies with the RF. After the feature selection, the final MLSE estimation model with the RF can be expressed as MLSE=*f*(MPDI, PR, elevation, sand, roughness, silt).

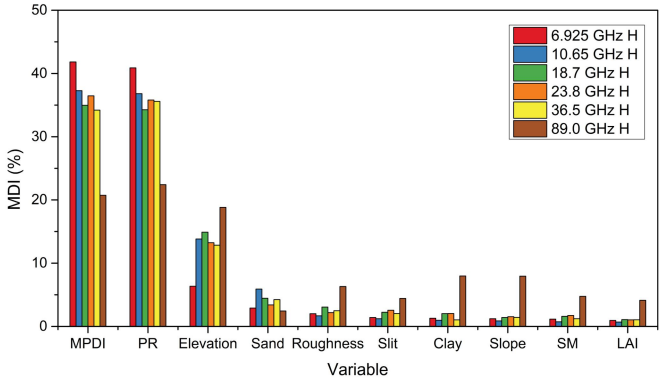


Fig. 3. The MDI% of ten variables at six AMSR-E frequencies.

Cloudy-sky MLSE can be estimated using the constructed model above. However, given that the RF requires sufficient observation samples for training complex machine learning model and has poor spatiotemporal migration capability since it cannot extrapolate target values outside the range of the training data [61], we selected clear-sky MLSEs on specific date t and its adjacent time-series data before and after 1 day (date t-1 day to t+1 day) as a suite of the datasets to train the model and then predict the cloudy-sky MLSE on t day using the above-selected optimal impact variables. There are two main reasons for choosing three days as the time window for the data training: 1) the MLSE is time-dependent and changes obviously over larger temporal windows; and 2) the selected time size can guarantee the data training volume and save training time to a certain degree. However, it should be noted that each model is trained based on the clear-sky MLSE

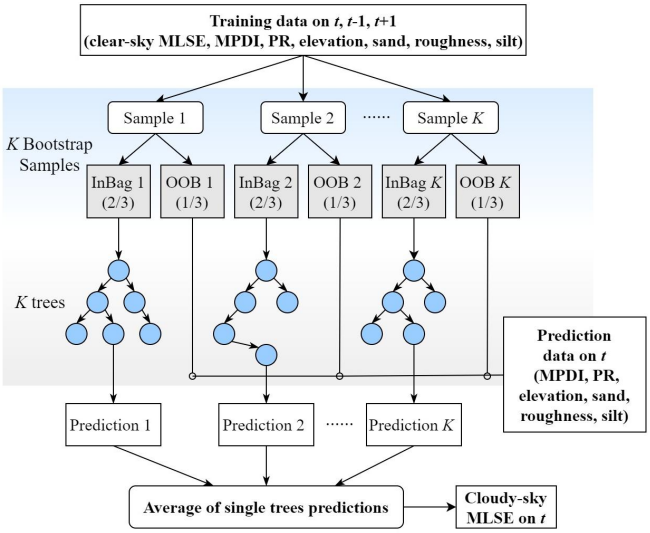


Fig. 4. The training and prediction of cloudy-sky MLSE with RF model.

obtained on a specific day; thus, all generated RF models are specific to the date to be analyzed. The specific process of training and predicting the cloudy-sky MLSEs using the RF model is given in Fig. 4. After the cloudy-sky MLSEs were estimated, we performed quality control and checked the instantaneous microwave emissivity datasets by masking the rain-covered, snow-covered, and water bodies-covered regions and excluding the emissivity greater than 1 and less than 0 in the end.

C. Accuracy Evaluation

Due to the lack of MLSE field measurements, Qiu_MLSE is used as a reference dataset to verify our MLSE dataset under clear-sky conditions. For convenient comparison, three widely used metrics are used to quantify its accuracy, including the root-mean-square error (RMSE), relative RMSE (RRMSE), and coefficient of determination (R^2). The RMSE and RRMSE are adopted to evaluate the consistency between the predicted MLSE ($MLSE_{\text{pre}}$) and Qiu_MLSE ($MLSE_{\text{ref}}$), and R^2 is used to characterize the similarity between the predicted MLSE ($MLSE_{\text{pre}}$) and Qiu_MLSE ($MLSE_{\text{ref}}$). In general, in an ideal circumstance, an RMSE close to 0 indicates the predicted value is close to the actual situation. Additionally, an RRMSE well below 0.5 denotes that the used method is more reliable, and an R^2 close to 1 indicates the predicted image detail is more similar to the actual image. Three indices are calculated as follows, respectively:

$$RMSE = [\frac{1}{n}\sum_{i=1}^{n}(MLSE_{\text{pre}} - MLSE_{\text{ref}})^2]^{1/2} \tag{6}$$

$$RRMSE = \frac{RMSE}{\frac{1}{n}\sum_{i=1}^{n} MLSE_{\text{ref}}} \tag{7}$$

$$R^2 = 1 - (\sum_{i=1}^{n}(MLSE_{\text{pre}} - \overline{MLSE_{\text{pre}}})^2 / \sum_{i=1}^{n}(MLSE_{\text{ref}} - \overline{MLSE_{\text{ref}}})^2) \tag{8}$$

where n is the number of pixels in MLSE image; $MLSE_{\text{pre}}$ is the predicted MLSE image; $MLSE_{\text{ref}}$ refers to the referenced Qiu_MLSE product; $\overline{MLSE_{\text{pre}}}$ is the mean of predicted MLSE; $\overline{MLSE_{\text{ref}}}$ is the mean of Qiu_MLSE.

TABLE II
SPATIAL RESOLUTIONS AND ACCURACIES OF SIX REANALYSIS SKIN TEMPERATURES

	ERA5	ERA-Interim	JRA55	MERRA2	FNL	GFS
Resolution	0.25°	0.75°	1.25°	0.25°	1.0°	0.5°
RMSE	5.73	5.85	16.83	7.12	10.88	10.48
RRMSE	0.55	0.56	1.61	0.68	1.04	1.01
R^2	0.88	0.88	0.49	0.80	0.82	0.74

IV. RESULTS

A. Impacts of Input Datasets on Clear-sky MLSE Inversion

The retrieval accuracy of clear-sky MLSE plays a key role in cloudy-sky MLSE estimation because the proposed method requires accurate clear-sky MLSE data as training samples to construct the estimated model. From Eq. (1), we perceive that the deviation of input datasets such as BTs and LSTs acquired by the sensors and atmospheric profiles would produce large errors for the MLSE calculation. Thus, it is essential to point out the impacts of input data on clear-sky MLSE estimates. Because the radiometric noise in the AMSR-E BT is less than 1.1 K [50] and induces small uncertainties in the retrieved surface emissivity [2], the MLSE retrieval errors caused by the uncertainties of LSTs and atmospheric profiles are merely discussed in this section.

The uncertainty in the LST is a leading source of MLSE error, especially at the microwave frequencies below 19 GHz, since the microwave radiation is more sensitive to the surface than to the atmosphere [40]. As we used the MODIS LST, which is of better accuracy, to calculate the MLSE under clear skies, the produced MLSE errors due to it are not discussed here. However, in previous studies, some MLSE estimations used the reanalysis skin temperatures as the LST to calculate the MLSEs [34, 39]. Hence, we mainly investigated the impacts of multiple reanalysis skin temperatures on MLSE estimates. In this study, the selected reanalysis skin temperatures derive from the ERA5, ERA-Interim, JRA-55, MERRA2, FNL, and GFS. Taking the MODIS LST as true values, Table II first compares the differences among the six types of reanalysis skin temperatures under clear skies at 23 sounding stations in 2008. As shown in Table II, the RMSE differences between six reanalysis skin temperatures and MODIS LST are all greater than 5.73 K and the RRMSE differences are all greater than 0.5 K, indicating they all have relatively poor accuracies. In contrast, the ERA5 data has the optimal performance, followed by the ERA-Interim product, MERRA2, NCEP/GFS, JRA55, and NCEP/FNL. Moreover, using the clear-sky MLSEs retrieved with the UWYO profiles, AMSR-E BTs, and MODIS LSTs at 23 sounding stations as true MLSE values, Fig. 5 also investigates the impacts of six reanalysis skin temperature products on the clear-sky MLSE retrieval. The impact analysis in Fig. 5 shows that the LST uncertainty is a pivotal factor impacting the estimate accuracy of MLSE. Regarding the H polarization, the RMSEs of clear-sky MLSE calculated caused by the ERA5 skin temperature vary from 0.0137 at 6.9 GHz to 0.0212 at 89 GHz,

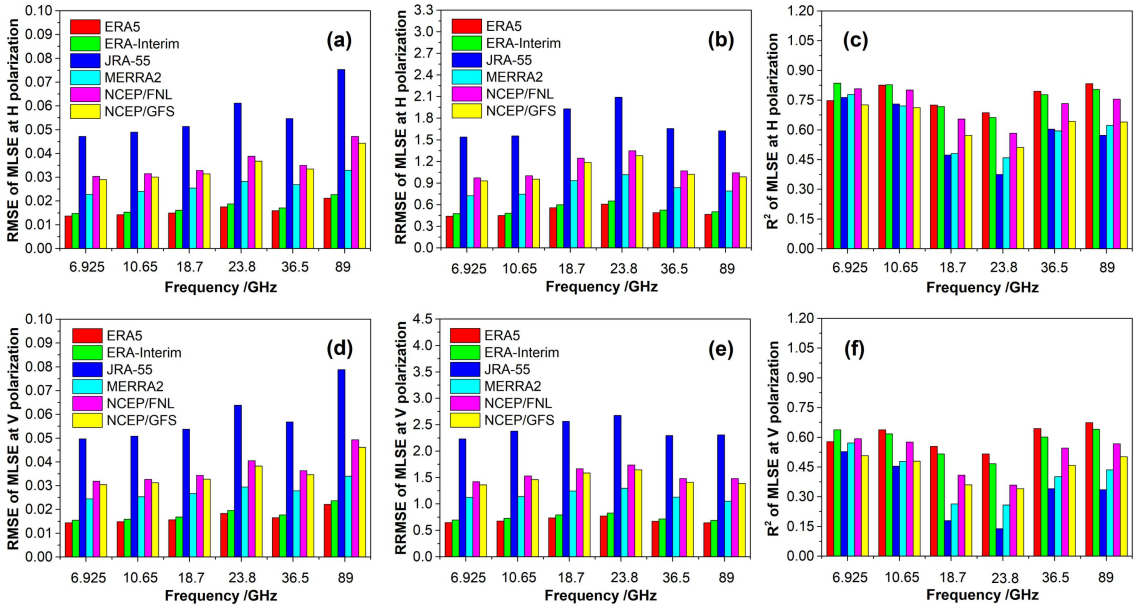


Fig. 5. Impacts of six reanalysis skin temperatures on the calculation of clear-sky MLSE for six AMSR-E frequencies. (a) RMSE change at H polarization; (b) RRMSE change at H polarization; (c) R^2 change at H polarization; (d) RMSE change at V polarization; (e) RRMSE change at V polarization; (f) R^2 change at V polarization.

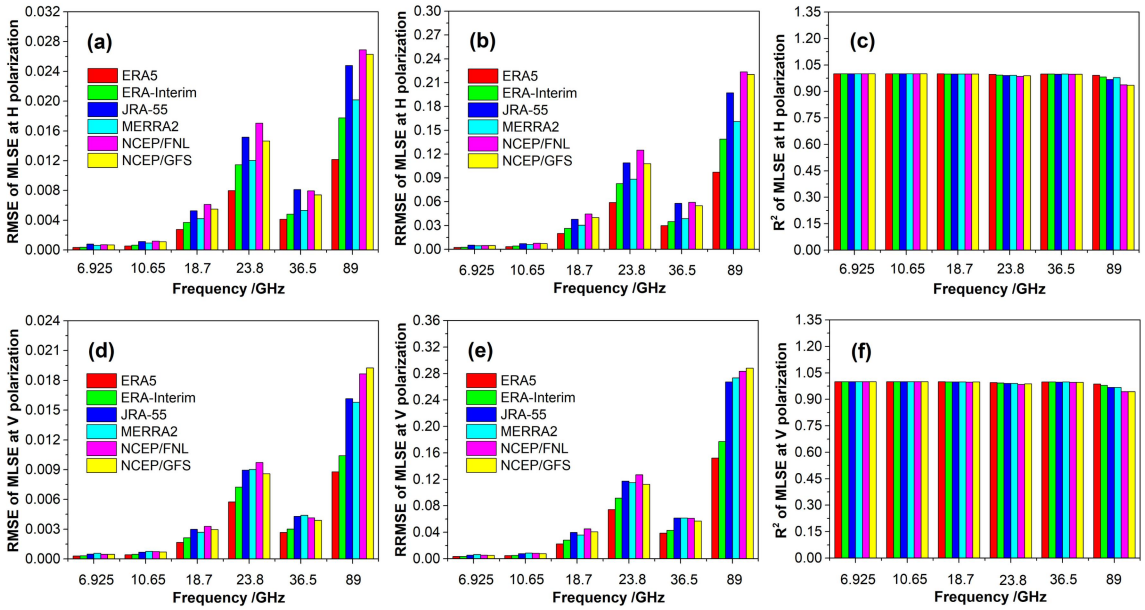


Fig. 6. Impacts of six reanalysis profiles on the calculation of clear-sky MLSE for six AMSR-E frequencies. (a) RMSE change at H polarization; (b) RRMSE change at H polarization; (c) R^2 change at H polarization; (d) RMSE change at V polarization; (e) RRMSE change at V polarization; (f) R^2 change at V polarization.

and its RRMSEs vary from 0.442 at 6.9 GHz to 0.607 at 23.8 GHz. Obviously, the RMSEs in the emissivity retrieval may be as high as 0.0137 or more for a 5.73 K error of ERA5 skin temperature. Therefore, it can be seen that the previous studies that used some reanalysis skin temperatures to perform the acquisition of MLSEs are flawed.

In addition to the impact of LST uncertainty, the uncertainty in atmospheric profiles is another key factor causing MLSE error. Fig. 6 presents the assessment results of six reanalysis atmospheric profiles (i.e., the ERA5, ERA-Interim, JRA-55, MERRA2, NCEP/FNL, and NCEP/GFS) in estimating MLSE under clear skies. The clear-sky MLSEs retrieved with the UWYO profiles, AMSR-E BTs, and MODIS LSTs at the same stations are still used as references. Based on the RMSE, RRMSE, and R^2 statistics of the six reanalysis profile products in calculating clear-sky MLSEs (see Fig. 6), we discover that the retrieved MLSEs with six reanalysis profiles at frequencies with higher atmospheric attenuation present the larger errors (e.g., 23.8 and 89 GHz). In particular, RMSEs of retrieved MLSEs using six reanalysis profiles are quite large at 89 GHz (the RMSEs of all profiles are greater than 0.009 at this frequency). The specific reasons for the poor performance can be attributed to two aspects. First, these two frequencies tend to have a strong sensitivity to WV or CLW parameters. Second, as compared with the UWYO profile, all reanalysis profiles have lesser pressure levels, so the calculated three

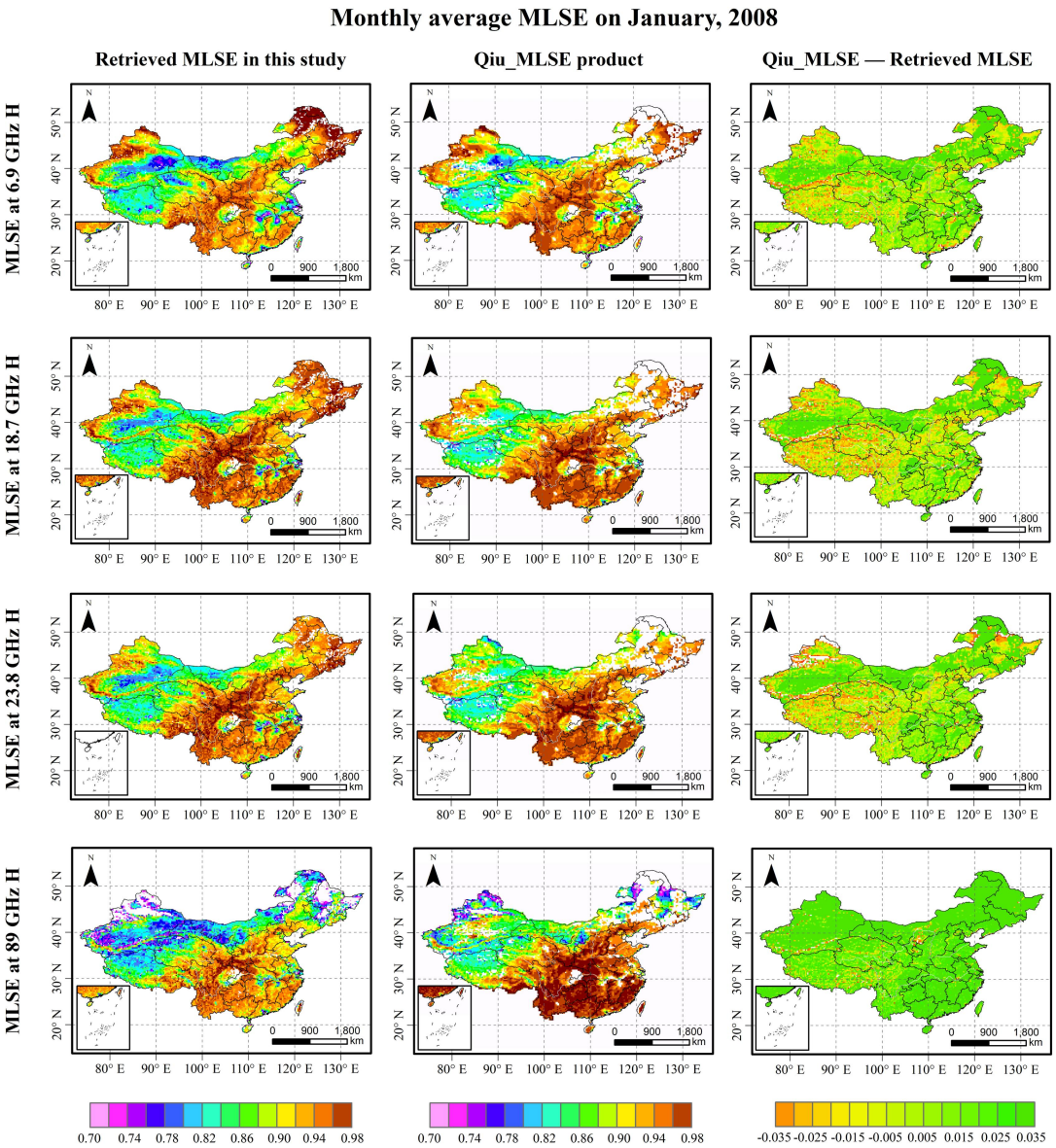


Fig. 7. The spatial distributions of monthly average clear-sky MLSEs derived from us and Qiu et al. in January 2008 for four AMSR-E frequencies. The corresponding difference maps between two datasets are also given.

atmospheric parameters (T_u, T_d, and τ) perform poorly for all profile types. In a physical sense the high air attenuation obscures the surface from view, which makes it more difficult to measure surface emissivity accurately. However, as a whole, the ERA5 profile generates the best performance. For the H polarization, its RMSEs vary from 0.0003 at 6.9 GHz to 0.012 at 89 GHz, and the RRMSEs vary from 0.002 at 6.9 GHz to 0.097 at 89 GHz. The uncertainties produced by the errors of the ERA-Interim, MERRA2, JRA-55, GFS, and FNL products all present gradually increasing trends. This indicates that the ERA5 has a desirable potential in terms of revealing actual atmospheric conditions and greatly improving the accuracies of clear-sky MLSEs estimates. This has been demonstrated in our previous research [53]. However, we observe that the impacts of the errors from profiles on the MLSE estimates are lower. An RMSE of 12.87 K in the T_d of NCEP/FNL produces an RMSE of 0.027 for the MLSE at 89 GHz. Similar results were also explored by Hu et al. [34], who found that the MLSEs presented a 1-2% error to 10% bias of the temperature profile. Some scholars also indicated that a 25% change in the water vapor would lead to a global mean change in the MLSE of 0.0016 at 6.9 GHz and 0.03 at 89.0 GHz [59, 60].

B. Performance of Clear-skies MLSE Retrieval

Because the estimated clear-sky MLSEs lack ground-truth measurements that can be validated, we indirectly compared the clear-sky MLSE estimated by us with the existing MLSE product, i.e., Qiu_MLSE, from the spatiotemporal distribution. Qiu_MLSE is mainly calculated under clear skies and applies the same BT and similar LST product as the clear-sky MLSE retrieved in this paper; thus, it is meaningful to perform the contrast with this product.

The AMSR-E BTs have some missing values due to their scanning swath. For convenient comparison, we calculated the monthly average MLSE maps with the valid pixels by filtering the cloudy and empty pixels. Fig. 7 presents the distribution

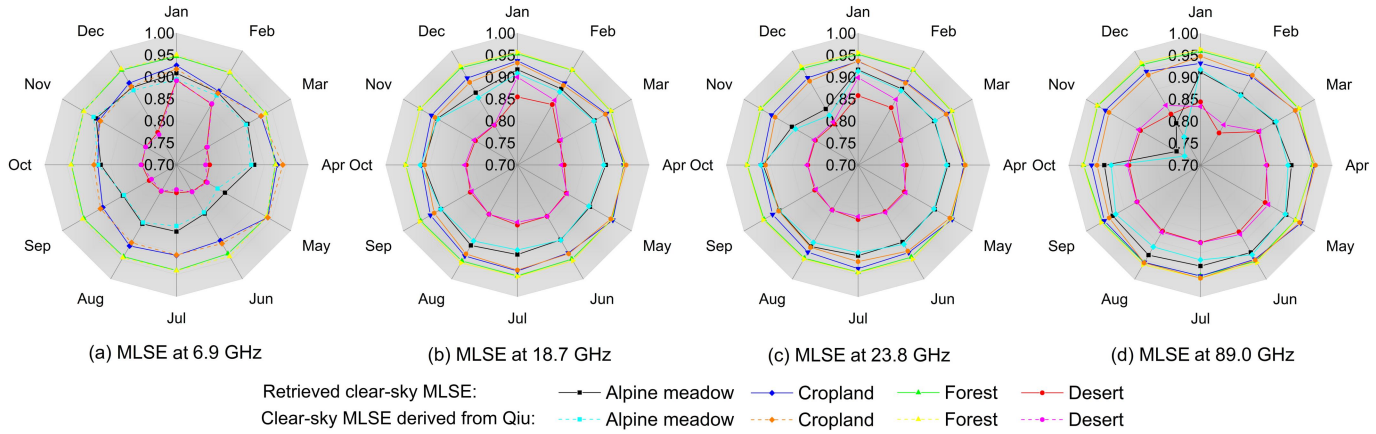


Fig. 8. The time-series differences between the retrieved MLSE and Qiu_MLSE at different months and sites. (a) 6.9 GHz, (b) 18.7 GHz, (c) 23.8 GHz, (d) 89 GHz.

maps of monthly average MLSEs in January 2008 at four AMSR-E frequencies (i.e., 6.9, 18.7, 23.8, and 89 GHz). From Fig. 7, we see the spatial distributions of monthly average MLSEs are highly consistent between the two MLSE datasets over China. They all display expected spatial structures that are closely associated with changes in frequencies, land-cover types, and topography. This further confirms the conclusion that the atmospheric profile has little influence on surface emissivity estimates. In addition, we note that the calculated monthly average MLSEs obviously increase with increasing frequency overall, but the high-latitude areas have some low MLSEs at 89 GHz in Fig. 7. This is since the continuous snowfall event in the winter presents a pronounced scattering impact on the MLSE calculations at higher frequencies. In terms of spatial distribution, the higher MLSEs are mainly concentrated in the Loess Plateau, the Qinling mountainous regions, the YunGui Plateau, and the southeast regions of the study area, in which the surfaces are covered mostly by forests, having a positive response to the MLSE at this frequency. In contrast, low MLSE values are mostly found in the northern Qinghai-Tibet Plateau and the northern Inner Mongolian Plateau, where the land surfaces are covered mostly by sparse vegetation and deserts. In addition, some rivers and coastal regions also exhibit low to moderate MLSEs mainly due to the prominent scattering derived from the water body and river. Based on the distribution maps of differences between these two MLSE datasets, we further validate that the retrieved MLSEs are in line with Qiu_MLSE, with differences ranging between -0.035−0.035 in January 2008 for the four selected frequencies. However, the clear-sky MLSEs estimated have underestimates in some mountainous regions (in particular, the Qinghai-Tibet Plateau) at 6.9, 18.7, and 23.8 GHz and present an overestimate at 89 GHz as a whole. This maybe is associated with the vertical stratification features of atmospheric profile and microwave scattering. On the whole, the clear-sky MLSEs estimated in this study have a respectable precision and good spatial distribution performance.

According to the statistical analysis of the monthly average MLSEs from January to December 2008 at four land-cover types: alpine meadow, desert, forest, and cropland, Fig. 8 further compares the time-series changes of the retrieved MLSE in this paper and Qiu_MLSE at different months and sites. These four kinds of land-cover types present different sensitivities to MLSE changes, which can reflect the overall trend of MLSE changes under actual natural conditions. From the MLSE temporal changes, we discover that the retrieved MLSEs under clear skies are accordant with Qiu_MLSE each month. There is a small MLSE change over a short period of time, and the possible impact of cloud covers on MLSEs is mild. Especially for the forest and croplands, the two datasets in these regions are more coincident, even if the water vapor emission and heat energy transmission from the vegetation canopy or the transpiration accelerate the formation of clouds and rain. Since the SM and vegetation cover status in the Gobi deserts and forests do not have obvious changes, the MLSEs in these regions are more stable with little seasonal variations. However, an abnormal phenomenon is perceived in the desert region in which its MLSE values at all frequencies increased suddenly in January and February 2008. This result is likely because the snow cover in the Takla Makan Desert at that time generated a typical scattering effect, and its MLSEs increased with the decreasing frequency. Nevertheless, for the alpine meadow and croplands, due to the growth change of grassland and the irrigation period of crops, an evident seasonal change trend is observed on the MLSE. Based on the above analysis, it is not difficult to imagine that these anomalous MLSEs are relevant to underlying surface variation and human activity.

C. Performance of the Cloudy-sky MLSE Estimates

It is essential to perform a systematic assessment of the RF model before the cloudy-sky MLSEs are estimated for further determining the applicability of the RF model. In this section, approximately two-thirds of clear-sky MLSEs and predicted factors from 31 December 2007 to 2 January 2008 were randomly selected for training and tuning the RF model, and the remaining one-third of the data pairs were used to evaluate the performance of the MLSE estimation model. Taking the estimated MLSE at 18.7 GHz as a sample, Fig. 9 exhibits the performances of the RF models in predicting clear-sky MLSEs at H polarization, V polarization, and polarization difference (V-H). As shown in Fig. 9, in these three RF models, most

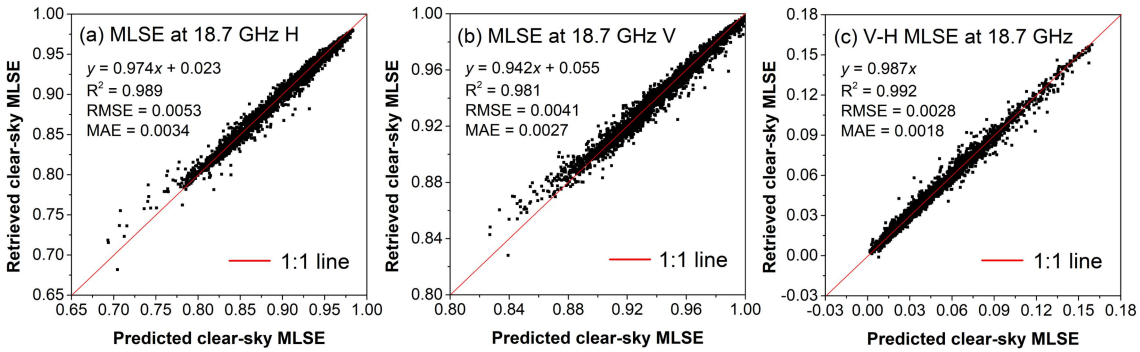


Fig. 9. Scatterplots of the predicted clear-sky MLSE against the retrieved clear-sky MLSE at (a) H polarization, (b) V polarization, and (c) polarization difference V-H.

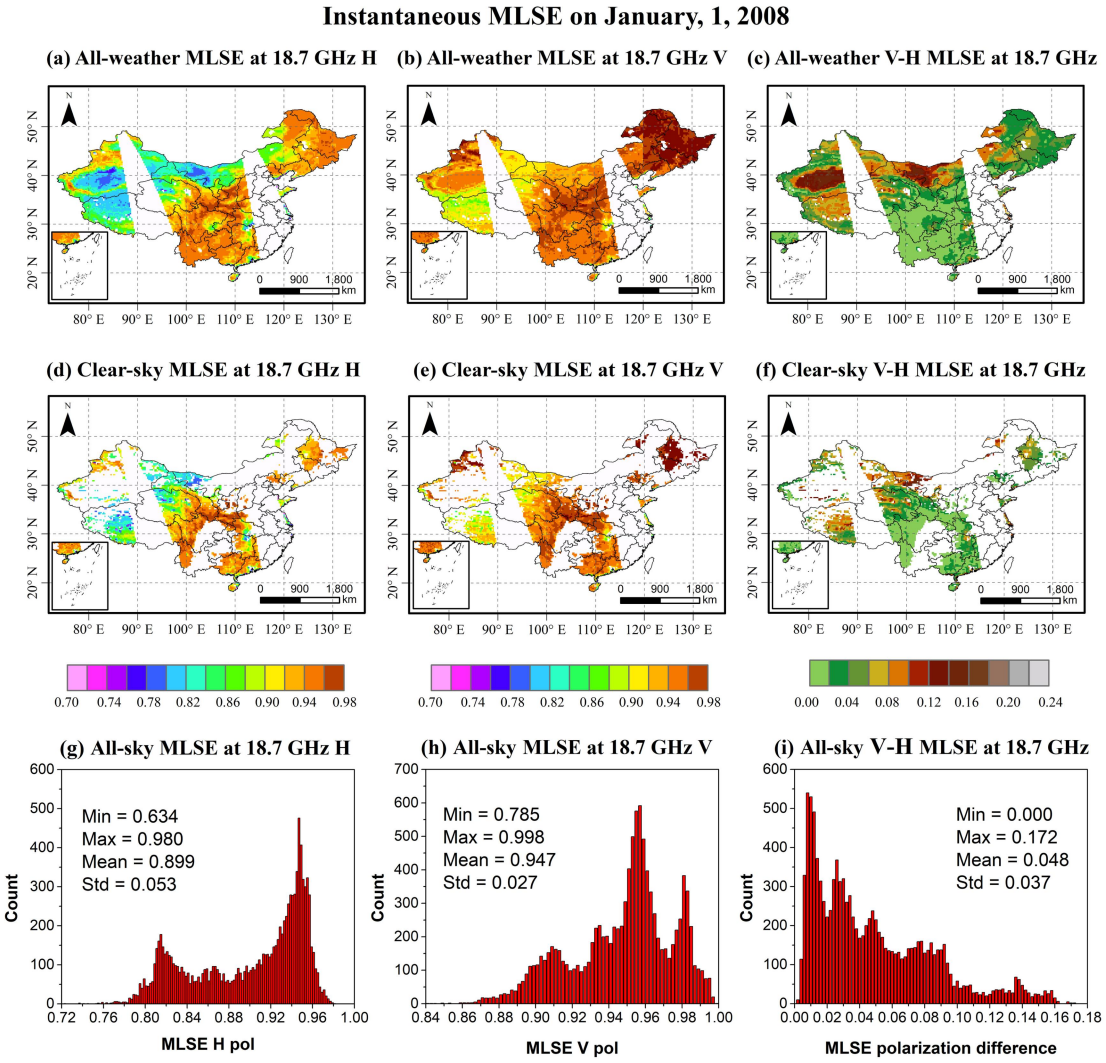


Fig. 10. The estimated all-weather MLSEs and the retrieved clear-sky MLSEs in China on 1 January 2008. The corresponding all-weather MLSE statistic are also given.

scatter points are concentrated near the 1:1 line, suggesting that these prediction models can obtain satisfactory accuracies in estimating instantaneous clear-sky MLSEs. In addition, they display better validation statistic results, with the R^2, RMSE, and MAE ranging from 0.981 to 0.992, 0.0028 to 0.0053, and 0.0018 to 0.0034, respectively. Nevertheless, we also perceive that the estimated clear-sky MLSEs at the H polarization are poorer than the MLSEs at the V polarization and polarization difference (V-H). This suggests that the relationship between the clear-sky MLSEs and selected auxiliary variables for the V polarization is stronger. In addition, a similar finding is also presented in other AMSR-E frequencies, where the RMSEs of

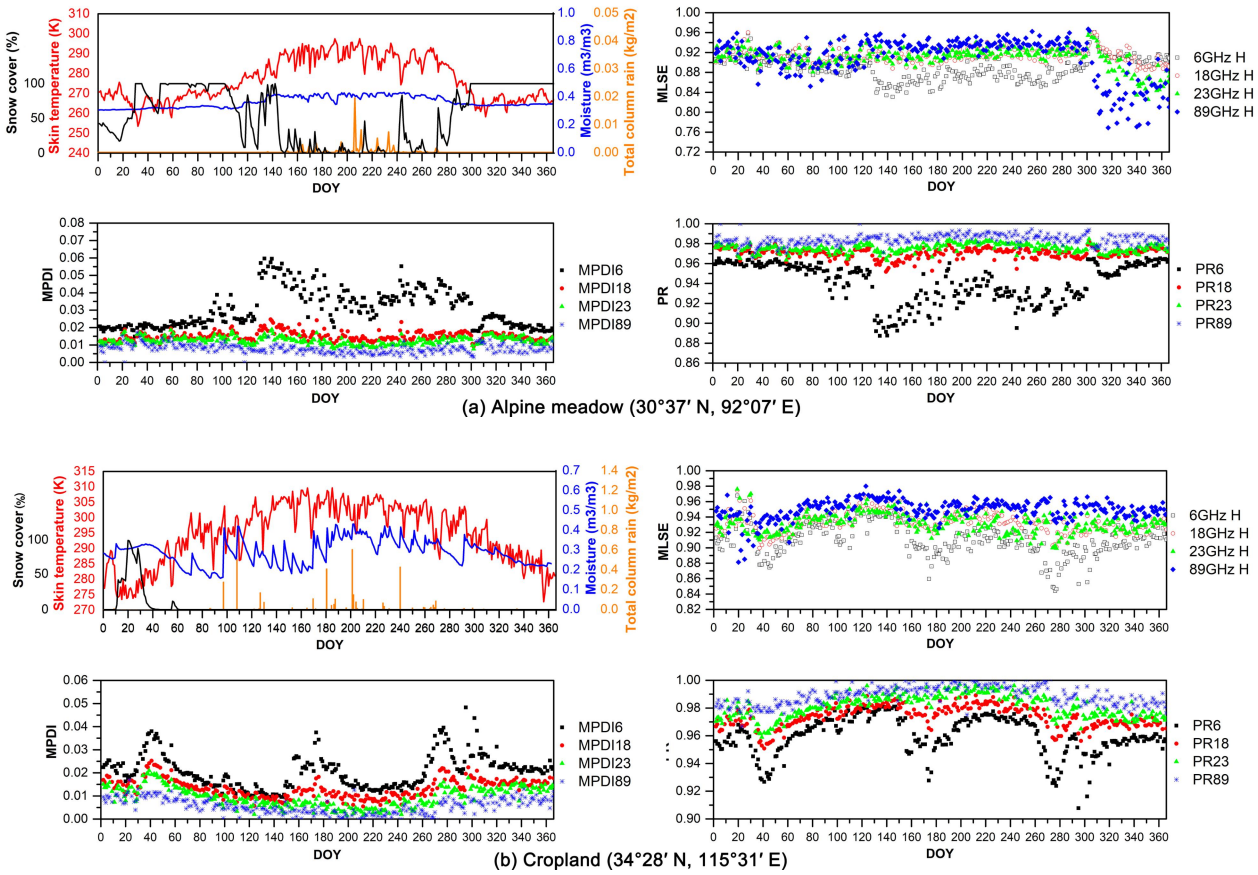


Fig. 11. Time-series of AMSR-E derived horizontally polarized MLSEs, MPDI, and PR at four AMSR-E frequencies at four sites: (a) alpine meadow (b) cropland. The corresponding curves of snow cover (black line), total precipitation value (Orange line), skin temperature (red line), and volumetric soil water (blue line) observed at four land-cover types in China for 2008 are also shown.

the estimated MLSEs at the V polarization are all less than 0.005. This is likely because the MLSE at the H polarization is more sensitive to surface variables such as the soil roughness, SM, and vegetation water content. However, as a whole, the estimated clear-sky MLSEs with the RF model have higher consistency with the retrieved clear-sky MLSEs, and there are few differences. This finding demonstrates the selected factors could better estimate MLSEs, and the RF has a good potential to estimate cloudy-sky MLSE datasets over a wide range of spatiotemporal scales.

After extending the relationship between clear-sky MLSEs and related impact variables to cloudy conditions, Fig. 10 also displays a case for the prediction of all-weather MLSEs on 1 January 2008. For comparison purposes, the clear-sky MLSEs derived from the actual satellite instantaneous observations are displayed. From Fig. 10, except for the impact of the AMSR-E running orbit, 46.5% of the China region is covered by clouds, and our strategy effectively predicts all-weather MLSE values covering 97.3% of the total study area. This suggests that the proposed method can successfully recover the image gaps affected by cloud contamination as compared to the clear-sky MLSEs. There is a relatively clear image texture even if the MLSEs transition from the edge regions of the cloud cluster to the neighboring cloud-free regions. The spatial distributions of estimated MLSEs in cloudless areas are more consistent with the actual MLSEs. However, due to the impacts of snowfall or water bodies, some obvious image gaps are also observed in lakes and rivers. Moreover, the cloudy-sky MLSE estimation results present a good distribution that is closely associated with land cover status. In the vegetated regions, the surface emissivity reacts to the vegetation water content with a high MLSE value; however, in the semiarid regions, the emissivity polarization difference is relatively sensitive to the variability in SM, and the emissivity is smaller. Meanwhile, from Fig 10(g)-(i), we note that the cloudy-sky MLSEs estimated at the V polarization are generally greater than the cloudy-sky MLSEs estimated at the H polarization. There are some more concentrated MLSEs with a standard deviation (STD) of 0.027. This result is in line with the MLSE change in the actual situation and further demonstrates the dependability of the RF model used in this study.

However, it is worth noting that, in this study, we merely presented the spatial distribution of cloudy-sky MLSEs and did not validate its accuracy because of the unavailability of real surface emissivity products under cloudy-sky conditions. In future research, for further verifying the performance of the estimated cloudy-sky MLSE, using some auxiliary parameters related to the MLSE changes (e.g., SM, vegetation phenology) to analyze the spatiotemporal correlation is an indirect way.

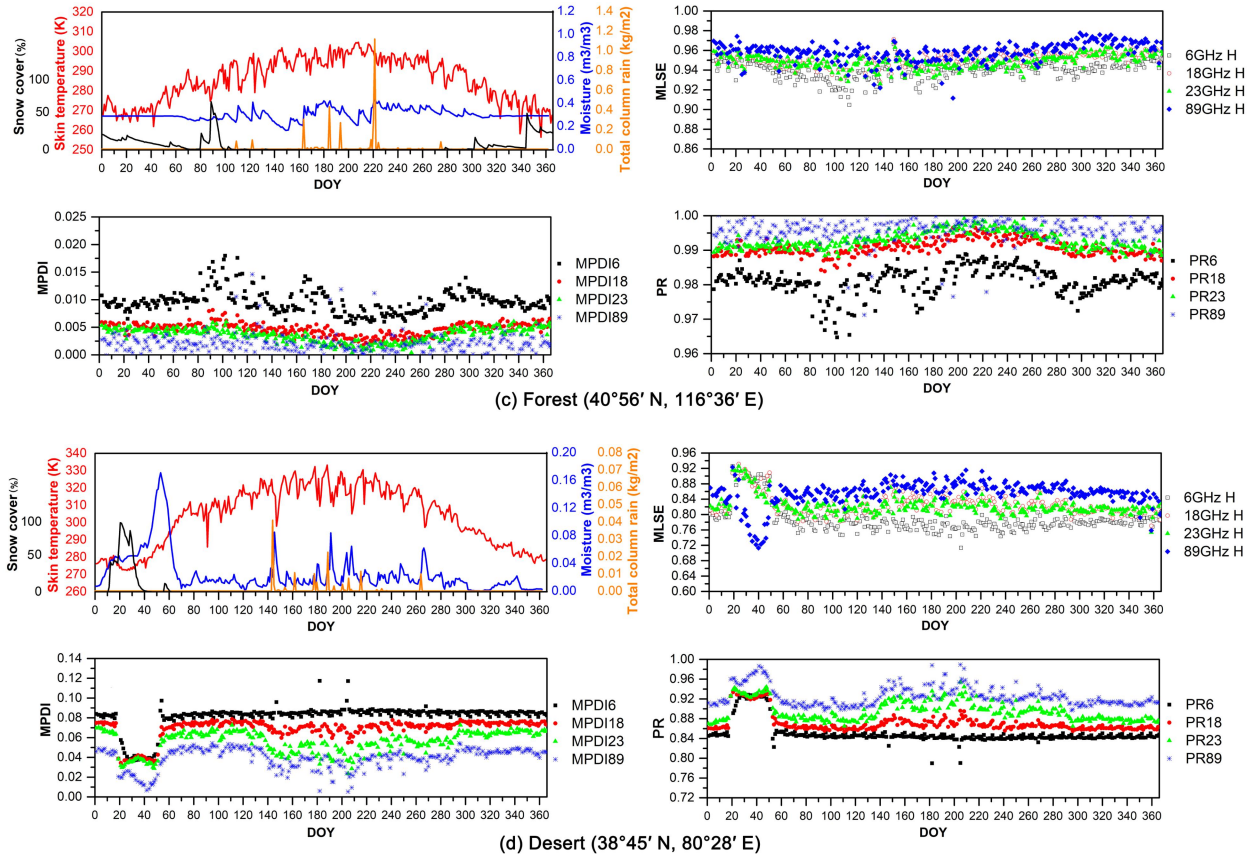


Fig. 11. Time-series of AMSR-E derived horizontally polarized MLSEs, MPDI, and PR at four AMSR-E frequencies at four sites: (c) forest (d) desert. The corresponding curves of snow cover (black line), total precipitation value (Orange line), skin temperature (red line), and volumetric soil water (blue line) observed at four land-cover types in China for 2008 are also shown.

V. Discussion

To provide more information regarding the estimates of MLSEs, there are still some points that need to be discussed, such as the potential impact factors of MLSE changes, the sensitivity of cloud properties to the MLSE, the comparison between the TIR-derived emissivity with clear-sky MLSE, and the limitations of this method. This section will discuss them.

A. The Possible Impact Factors of MLSE Retrieval

Previous surface experiments have shown that the MLSE varies considerably with changes in the SM and temperature [41, 67]. Precipitation and snow cover also greatly influence the MLSE since they can alter the land surface water potential and its dielectric properties [58]. Hence, in this section, taking the skin temperature, SM, snow cover, and total precipitation derived from the ERA5-Land hourly data as four key impact factors, we investigated the predominant factors affecting the MLSE at various times and regions. To reduce the differences among impact factors and MLSE in spatiotemporal resolutions, all impact factors have been adjusted to a 0.25° resolution. And then, the annual change curves of ERA5 skin temperature, volumetric soil water, snow cover, and total precipitation are respectively extracted at four sites with various land-cover types. Meanwhile, the corresponding time series of AMSR-E MLSEs at frequencies of 6.9, 18.7, 23.8, and 89 GHz and their PR and MPDI indices for 2008 are also given.

As shown in Fig. 11(a)-(d), the estimated MLSEs present significant variability with fluctuations in these four impacting factors at the four specific sites in the given year. The MLSEs exhibit apparent sensitivities to four selected parameters, and the PR index is negatively correlated with the MPDI and has a positive relationship with the MLSEs at the corresponding frequencies. However, it is noted that skin temperature does not display a strong response to the MLSE. These emissivity variations are not correlated with the fluctuations of soil skin temperatures, even if the skin temperature plays an essential role in controlling the land surface heat and evapotranspiration of soil/vegetation water content. Moreover, we observe that the predominant factors influencing the changes in MLSE are obviously different among the four land-cover types, and the contribution intensity of different components of the surface to the MLSE is highly frequency dependent. For the meadow and cropland, the MLSEs at 6.9 GHz respond significantly to soil water content due to a strong correlation between the MLSE and change in the SM. In particular, in the cropland site, the MLSEs and soil water display relatively consistent seasonal change trends due to the on-farm irrigation during the crop growing seasons (early spring, summer, and mid-autumn). The alpine meadow site also presents a sudden change from June to September 2008 because of the regular rainy season on the Qinghai-Tibet Plateau. Meanwhile, the MLSEs at the selected four frequencies become very low due to the unstoppable snowfall in early spring and winter, suggesting that the snow

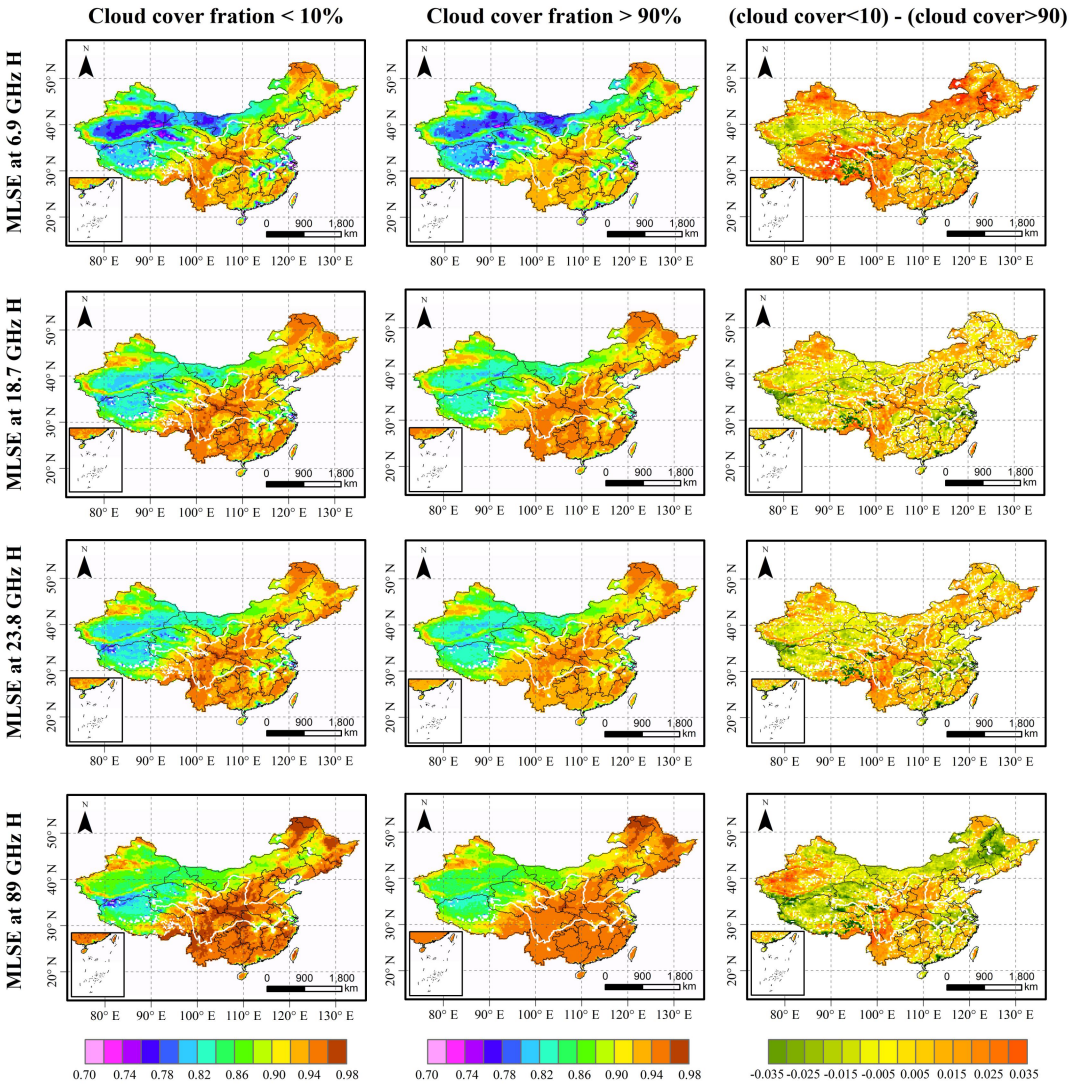


Fig. 12. The yearly average distributions of MLSEs under clear skies (total cloud cover < 10%) and heavy cloudy skies (total cloud cover > 90%) in 2008 for four AMSR-E frequencies. The maps of differences between the two MLSEs are also given.

accumulation event acts as a perturbation factor affecting the seasonal variation in the MLSE. However, for the remaining two sites, their MLSE amplitudes are relatively stable for four AMSR-E frequencies throughout the year. A slight fluctuation trend is observed in the woodlands in the spring when the vegetation canopy is relatively sparse, and the rainfall ratio is increasingly frequent. However, a peculiar phenomenon of MLSE in the desert site was found in January 2008 when an evident trend of a sudden increase and then a decrease in the MLSE was observed. This pattern is likely because, in the transition season from winter to spring, when the snow begins to melt and bare soil becomes wet, the MLSE is affected by the SM and disturbed by the snowfall, and its fluctuation is relatively large. As a consequence, comparing the MLSE retrieval result with previous snowfall events is also an indirect way to show the MLSE on rainfall-affected surfaces.

B. The Sensibility of Cloud Property to MLSE Estimation

Aires et al. [23] employed the neural network approach to retrieve surface emissivity simultaneously based on the less difference in MLSE between clear and cloudy skies. Hu et al. [34] noted that the CLW had a negligible impact on MLSE calculation. Thus, the estimation of cloudy-sky MLSE in this study is mainly based on the hypothesis that the relationships between clear-sky MLSE and related impact variables could be extended to cloudy-sky conditions. Nevertheless, to further verify this hypothesis, we discussed the sensitivity of cloud cover status to the MLSE. To obtain information on the cloud properties that are consistent with AMSR-E observations in both space and time, the total cloud cover dataset derived from the ERA5 hourly data on single levels is used in this section.

Fig. 12 first compares the yearly average MLSE distribution maps under clear sky weather (total cloud cover < 10%) and heavy cloudy sky weather (total cloud cover > 90%) during 2008. We found that the distribution of MLSE in China and the mean values under these two weather conditions do not change significantly for specific microwave frequencies. This indicates that the possible impact of clouds on MLSE is mild. From the point of view of contrastive analysis, it further demonstrates that the cloudy-sky MLSE retrieval is feasible.

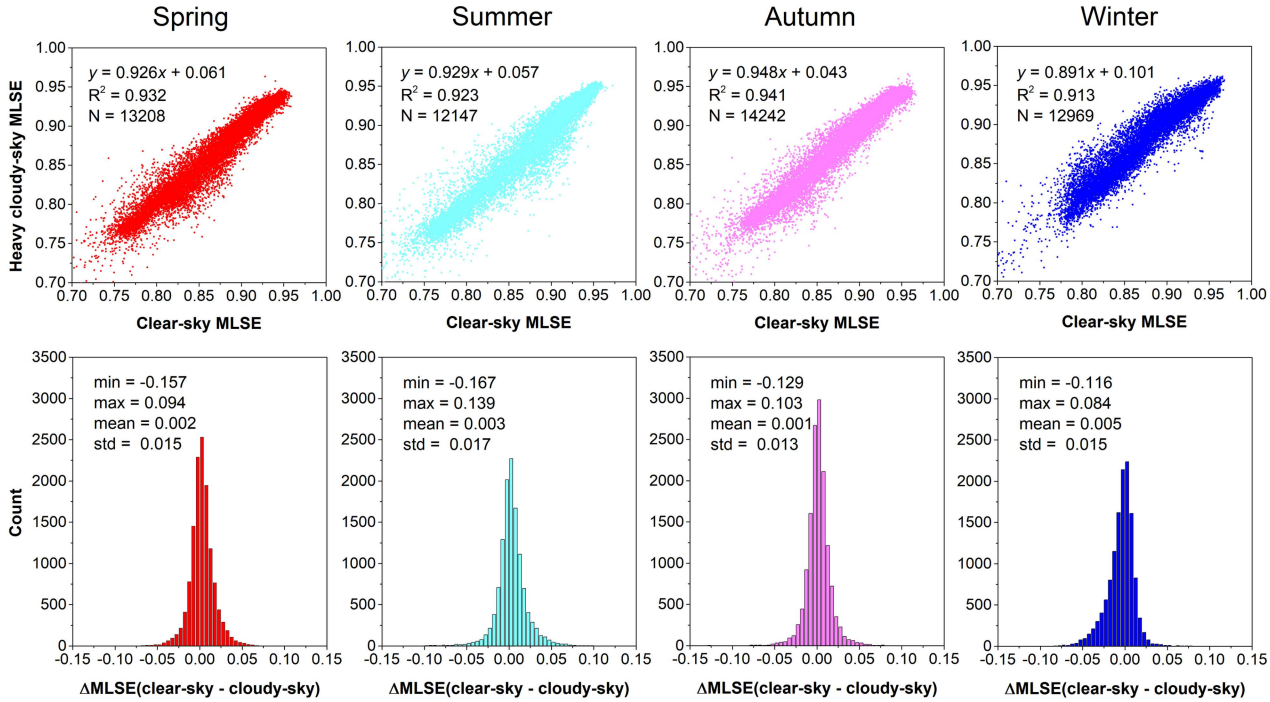


Fig. 13. The scatter diagrams between cloudy-sky MLSEs and clear-sky MLSEs and the difference value statistics of both in the four seasons of 2008.

Nevertheless, for 6.9 and 89 GHz frequencies, some apparent MLSE changes are displayed in the Northeast China Plain, the Taklimakan Desert, the Qinghai-Tibet Plateau, the Sichuan Basin, and the middle and lower reaches of the Yangtze River (see the right column in Fig. 12). This result may be due to the unique sensitivity of the MLSE to certain frequencies and the specific response of surface properties of these regions to the MLSE [68]. At the same time, their vulnerable landscape characteristics and particular climatic types are two important causes driving these findings. As we mentioned previously, the seasonal change in land-cover types and rainy seasons will obviously alter the vegetation's biological property and soil water-holding capacity on the land surface and further impact the spatiotemporal variations in the MLSE, leading to some abnormal values in the end.

Taking the MLSE map at 6.9 GHz as a case study, Fig. 13 further presents the scatter diagrams between the cloudy-sky MLSEs and clear-sky MLSEs in the four seasons of 2008 to explain the impacts of cloud cover on the MLSE. Meanwhile, the difference value statistics between both are also given in the probability distribution figure (see the bottom column of Fig. 13). As we have seen from scatter diagrams, there is a good correlation between cloudy-sky MLSEs and clear-sky MLSEs for the four seasons of 2008 (R^2 values are all greater than 0.9), whereas their correlation in the summer is poorer than that in the other seasons owing to the frequent heavy cloud cover. Despite this, from the probability distribution figure, we found that the differences between cloudy-sky MLSEs and clear-sky MLSEs are generally smaller and concentrate in a range from -0.05 to 0.05. The differences in the mean values are less than 0.005, and the STDs are less than 0.017 for the four seasons. In general, the possible impact of clouds on the MLSE is also weak in different seasons.

C. Comparison between Thermal Infrared Emissivity with the MLSE

Similar to the MLSE, land-surface emissivity derived from the TIR band is also an important input variable for estimating surface biophysical parameters, such as LST, land-cover types, evapotranspiration, and PWV for a variety of research projects. However, owing to the obvious difference between the two in the physical mechanisms, there are different distributions and change trends. Thus, in this section, we will further reveal the changes of MLSE and compare the spatiotemporal difference between TIR-derived emissivity and PMW-derived emissivity. In consideration of the excellent temporal collocation between AMSR-E and MODIS on the Aqua, the TIR emissivity from the MODIS bands 31 and 32 are used.

By averaging the clear-sky TIR emissivities and microwave surface emissivities for each day in 2008, Fig. 14 displays the yearly average distribution maps of surface emissivities in 2008 for two MODIS bands and four AMSR-E frequencies. Because of much higher variations in the emissivities in the microwave range and the dependence of microwave BTs on surface roughness and structures, yearly average MLSEs at four AMSR-E frequencies all exhibit a larger variation range from 0.70 to 0.98, and their spatial distributions closely respond to vegetation coverage degree. Whereas the spectral emissivity characteristics of the MODIS TIR bands 31 and 32 (wavelength range is 10.5-12.5μm) are very high compared to MLSE. There is a smaller change range from 0.965 to 0.993 for these two MODIS TIR bands at various terrestrial land cover backgrounds. Meanwhile, their spatial distributions are closely related to land-cover types.

In addition, by making a statistic for the yearly average emissivities at 15 land-cover types, we further observe from Fig. 15 that the MLSEs at four frequencies dramatically vary

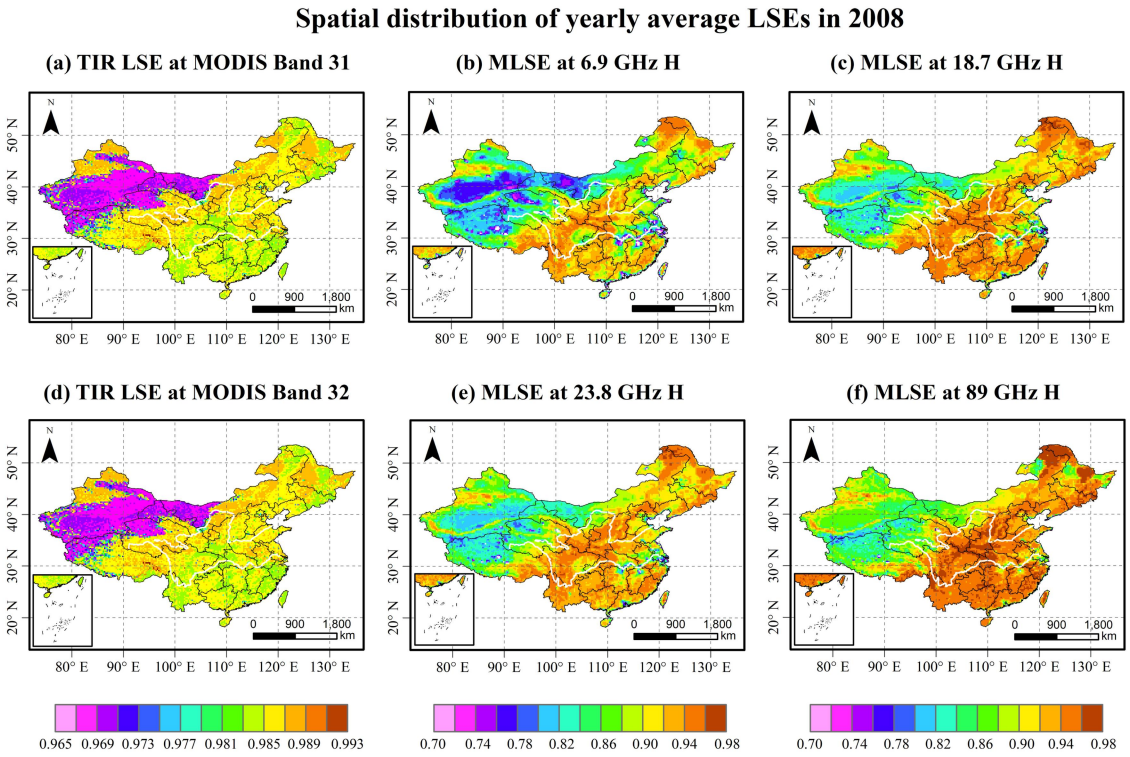


Fig. 14. The yearly average distributions of surface emissivities in 2008 for two MODIS bands and four AMSR-E frequencies.

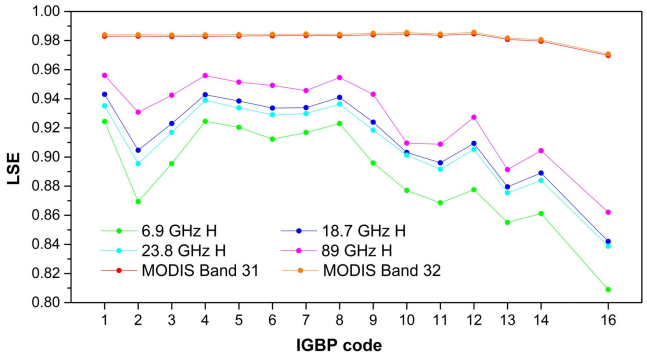


Fig. 15. The yearly average change trends of land surface emissivities at 15 land-cover types in 2008 for two MODIS bands and four AMSR-E frequencies.

by land-covers, while the TIR-derived emissivities have a steady trend. Taking 89 GHz as a case, for 15 surface types, the vegetation-covered regions (IGBP code 1, 3, 4, 5, and 6) display high emissivity values, followed by the grassland or cropland (IGBP code 10, 11, 12, and 14), and the urban area or barren regions (IGBP code 13 and 16). Under various land-cover types, there is a relatively steady emissivity change curve for the MODIS TIR bands 31 and 32. This finding has been proved by the existing spectral emissivity measurements of actual terrestrial materials, which indicate that the band average emissivities in MODIS bands 31 and 32 are relatively stable and known within about 0.01 for some land cover types, such as dense evergreen canopies, lake, snow surface, and most soils [69, 70]. Nevertheless, in the microwave frequency range, different from the TIR spectrum range, soil emissivities can rapidly vary from 0.6 for wet soil (about 30% volumetric soil moisture) to 0.9 for dry bare soil (about 8%) [71]. The emissivity of vegetation-covered surfaces varies obviously with the frequency and depends on the vegetation growth characteristics [72].

D. Limitations

Remotely sensed physical retrieval based on the RTE could simulate relatively accurate microwave surface emissivities, providing considerable convenience for building MLSE data. The RF has some good advantages over traditional regression methods for the accurate estimations of various atmospheric and geophysical parameters. As a result, combining satellite retrieval and machine learning model to realize the cloudy-sky MLSE retrieval has considerable superiority. In this paper, using each advantage of satellite observations and the RF algorithm, the proposed MLSE estimation method can derive cloudy-sky MLSEs at large spatiotemporal scales and present significant application potential for vegetation phenology monitoring, SM retrieval, atmospheric parameters derivation, and snow characteristic detection. However, some limitations are worth mentioning.

First, after the clear-sky MLSE is estimated, this study used the RF algorithm to generate cloudy-sky MLSEs. This process requires sufficient observation samples for the training model and will waste time. To increase the model training speed, we selected clear-sky MLSE datasets on date t and its adjacent time-series data before and after one day (date t-1 to t+1) as a suite of the data sets to train the model and then predicted the cloudy-sky MLSEs on t with the selected variables. This will lead to the generated RF models being specific to the day being analyzed. Moreover, the accuracy of the RF model is directly related to the regional distribution of training samples.

Therefore, the impact of the regional distribution of training samples on the RF model also needs to be considered.

Second, in addition to the effect of input data on the MLSE estimate, the quality of the method depends on other factors, including the frequency, polarization, sensor overpass time, incident angle, footprint, etc. [73]. Many previous studies also have indicated that the MLSEs generated by different sensors possess larger discrepancies, which are mainly caused by systematic errors and random errors in the retrieval process [20, 41]. At the same time, the systematic errors of MLSEs significantly vary in different regions. Therefore, if this new method can be applied to other sensor images and regions still requires additional research.

Third, a comprehensive assessment of the uncertainty in the MLSE estimate is a critical task for its wider applicability. However, the lack of ground truth measurements at a large scale makes validation difficult. In this study, we overcame this lack of ground truth data by investigating the consistency among the available microwave emissivity product from Qiu et al. since this product used similar input datasets. However, this comparison process was performed only under cloudless skies. For further evaluation, investigating its relations with some biophysical parameters, such as soil water content and vegetation phenology, both spatially and temporally, is another thinking that should be used in the future.

In addition, we should note that the derived MLSEs in this study only represent the effective surface emissivity due to the inconsistent physics–volume-averaged temperature from the PMW remote sensing versus skin temperature from the TIR data. In addition, in the field of view of AMSR-E detection, the pixel information is composed of different land-cover backgrounds. These key problems are inevitable because of the inherent nature of microwaves. Despite this, the acquired MLSE still presents good scientific significance for revealing the spatiotemporal changes in biogeochemistry, climatology, hydrology, and ecology sciences.

VI. Conclusion

The rapid development of remote sensing technology has produced a large number of algorithms for extracting MLSE from PMW observations. Nevertheless, to our knowledge, there has been limited research on estimating microwave emissivity under various weather conditions. In this study, a new method is proposed to estimate cloudy-sky MLSEs in China by combining each advantage of satellite observation and machine learning algorithm. The main conclusions are as follows.

(1) The accurate calculation of clear-sky MLSE plays a vital role in deriving the cloudy-sky MLSEs, which is subject to the precision of input datasets. This study selected the ERA5 reanalysis profile and MODIS LST to perform the acquisition of clear-sky MLSE, which is of more significant meaning. The LST accuracy has evident impacts on the uncertainty of the MLSE retrieval compared to that of atmospheric profiles. The calculated atmospheric parameters derived from the ERA5 are better than those output by the rest of the reanalysis profiles, and the estimated MLSE RMSEs obviously vary from 0.0003 at 6.9 GHz to 0.012 at 89 GHz.

(2) Overall, the spatiotemporal distributions of clear-sky MLSEs retrieved in this study are consistent with the product developed by Qiu et al. The MLSE distribution is closely associated with the changes in frequencies, surface types, and topography. The differences between the estimated monthly average MLSEs and Qiu_MLSE range from -0.035 to 0.035, and the error distributions present an obvious underestimate as a whole. Moreover, the seasonal cycle of clear-sky MLSEs estimated is also coincident with the Qiu_MLSE product at four kinds of land-cover backgrounds.

(3) The constructed RF model exhibits a desirable accuracy in predicting the clear-sky MLSE, along with an RMSE of 0.0028 in the 18.7 GHz H polarization. These models can be extended to the cloudy-sky status and successfully obtain an MLSE of over 97.3% for the satellite-detected land area on a typical cloudy day with a cloud fraction of 46.5%. Meanwhile, after filling image gaps affected by cloud-cover contamination, the proposed method generates a relatively clear and abundant MLSE distribution pattern compared to the clear-sky MLSE maps.

(4) Impact factor discussion of MLSE change, sensitivity analysis of cloud-cover properties on MLSE estimates, and the comparison between the TIR emissivity and MLSE suggest that the MLSE is modulated by many surface parameters, such as the LAI, SM, precipitation, snow cover, and land-cover types. However, it is less impacted by cloud cover conditions. This finding demonstrates that the proposed method is of evident potential in deriving cloudy-sky MLSEs over a global scale. However, some uncertainties still exist in this method, such as the error source, method applicability, and verification in different regions and sensors. Further research on the cloudy-sky MLSE estimation is still needed.

Acknowledgment

The authors would like to acknowledge the U.S. National Aeronautics and Space Administration (NASA) for the provision of the MODIS LST and the ASTER GDEM dataset (https://search.earthdata.nasa.gov/search); the National Snow and Ice Data Center (NSIDC) for the provision of AMSR-E brightness temperature and related datasets (https://nsidc.org/); the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing the ERA5 reanalysis data product (https://cds.climate.copernicus.eu/cdsapp#!/search?type=dataset). In addition, we also acknowledge the Computer Network Information Center, Chinese Academy of Sciences, China for the provision of global land surface emissivity data under clear skies status (https://datapid.cn/31253.11.sciencedb.113).

References

[1] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave remote sensing–Activeand passive: Microwave Remote Sensing Fundamentals and Radiometry*, 3 ed., p.^pp. 186–255, MA, USA: Addison-Wesley, Reading, 1981.

[2] C. Prigent, W. B. Rossow, and E. Matthews, “Microwave land surface emissivities estimated from SSM I observations,” *Journal*

of Geophysical Research Atmospheres, vol. 102, no. D18, pp. 21867-21890, 1997.

[3] A. S. Jones, and T. H. VonderHaar, "Retrieval of microwave surface emittance over land using coincident microwave and infrared satellite measurements," *Journal of Geophysical Research-Atmospheres,* vol. 102, no. D12, pp. 13609-13626, 1997.

[4] A. S. Jones, and T. H. Vonder Haar, "Passive microwave remote sensing of cloud liquid water over land regions," *Journal of Geophysical Research Atmospheres,* vol. 95, no. D10, pp. 16673-16683, 1990.

[5] S. J. English, "Estimation of temperature and humidity profile information from microwave radiances over different surface types," *Journal of Applied Meteorology,* vol. 38, no. 10, pp. 1526-1541, 1999.

[6] F. A. Furuzawa, H. Masunaga, and K. Nakamura, "Development of a land surface emissivity algorithm for use by microwave rain retrieval algorithms," *SPIE Asia-Pacific Remote Sensing,* vol. 8523, 8 November 2012, 2012.

[7] C. Birman, F. Karbou, and J. F. Mahfouf, "Daily rainfall detection and estimation over land using microwave surface emissivities," *Journal of Applied Meteorology and Climatology,* vol. 54, no. 4, pp. 880-895, 2015.

[8] D. B. Ji, J. C. Shi, C. Xiong *et al.*, "A total precipitable water retrieval method over land using the combination of passive microwave and optical remote sensing," *Remote Sensing of Environment,* vol. 191, pp. 313-327, 2017.

[9] J. P. Wigneron, A. Chanzy, J. C. Calvet *et al.*, "A simple algorithm to retrieve soil moisture and vegetation biomass using passive microwave measurements over crop fields," *Remote Sensing of Environment,* vol. 51, pp. 331-341, 1995.

[10] E. Njoku, T. Jackson, V. Lakshmi *et al.*, "Soil moisture retrieval from AMSR-E," *IEEE Transactions on Geoence & Remote Sensing,* vol. 41, no. 2, pp. 215-229, 2003.

[11] J. P. Wigneron, T. Jackson, P. O' Neill *et al.*, "Modelling the passive microwave signature from land surfaces: A review of recent results and application to the L-band SMOS & SMAP soil moisture retrieval algorithms," *Remote Sensing of Environment,* vol. 192, pp. 238-262, 2017.

[12] J. C. Shi, T. Jackson, J. Tao *et al.*, "Microwave vegetation indices for short vegetation covers from satellite passive microwave sensor AMSR-E," *Remote Sensing of Environment,* vol. 112, no. 12, pp. 4285-4300, 2008.

[13] L. J. Shi, Y. B. Qiu, and J. C. Shi, "Study of the microwave emissivity characteristics of vegetation over the northern hemisphere," *Spectroscopy and Spectral Analysis,* vol. 33, no. 5, pp. 1157-1162, 2013.

[14] R. Li, Y. P. Wang, J. H. Hu *et al.*, "Spatiotemporal variations of satellite microwave emissivity difference vegetation index in China under clear and cloudy skies," *Earth and Space Science,* vol. 7, no. e2020EA001145, pp. 1-15, 2020.

[15] F. Z. Weng, T. Zhu, and B. H. Yan, "Satellite data assimilation in numerical weather prediction models. Part II: Uses of rain affected microwave radiances for hurricane vortex analysis," *Journal of the Atmospheric Sciences,* vol. 64, no. 11, pp. 3910-3925, 2007.

[16] T. Holmes, M. Drusch, W. J. P. *et al.*, "A global simulation of microwave emission: Error structures based on output from ECMWF's operational integrated forecast system," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 46, no. 3, pp. 846-856, 2008.

[17] P. Bauer, G. Ohring, C. Kummerow *et al.*, "Assimilating satellite observations of clouds and precipitation into NWP models," *Bulletin of the American Meteorological Society,* vol. 92, no. 6, pp. ES25-ES28, 2011.

[18] F. Aires, C. Prigent, F. Bernardo *et al.*, "A tool to estimate land-surface emissivities at microwave frequencies (TELSEM) for use in numerical weather prediction," *Quarterly Journal of the Royal Meteorological Society,* vol. 137, no. 656, pp. 690-699, 2011.

[19] F. Z. Weng, B. H. Yan, and N. C. Grody, "A microwave land emissivity model," *Journal of Geophysical Research Atmospheres,* vol. 106, no. D17, pp. 20115-20123, 2001.

[20] S. Prakash, H. Norouzi, M. Azarderakhsh *et al.*, "Estimation of consistent global microwave land surface emissivity from AMSR-E and AMSR2 observations," *Journal of Applied Meteorology and Climatology,* vol. 57, no. 4, pp. 907-919, 2018.

[21] J. T. Pulliainen, J. Grandell, and M. T. Hallikainen, "HUT snow emission model and its applicability to snow water equivalent retrieval," *IEEE Transactions on Geoscience & Remote Sensing,* vol. 37, no. 3, pp. 1378-1390, 1999.

[22] K. S. Chen, T. D. Wu, L. Tsang *et al.*, "Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations," *IEEE Transactions on Geoence & Remote Sensing,* vol. 41, no. 1, pp. 90-101, 2003.

[23] F. Aires, C. Prigent, W. B. Rossow *et al.*, "A new neural network approach including first guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature, and emissivities over land from satellite microwave observations," *Journal of Geophysical Research Atmospheres,* vol. 106, no. D14, pp. 14887-14907, 2001.

[24] B. Ruston, F. Z. Weng, and B. H. Yan, "Use of a one-dimensional variational retrieval to diagnose estimates of infrared and microwave surface emissivity over land for atovs sounding instruments," *IEEE Transactions on Geoscience & Remote Sensing,* vol. 46, no. 2, pp. 393-402, 2008.

[25] S. A. Boukabara, K. Garrett, and C. Grassotti, "Dynamic inversion of global surface microwave emissivity using a 1DVAR approach," *Remote Sensing,* vol. 10, no. 5, pp. 1-18, 2018.

[26] F. Karbou, C. Prigent, L. Eymard *et al.*, "Microwave land emissivity calculations using AMSU measurements," *IEEE Transactions on Geoscience & Remote Sensing,* vol. 43, no. 5, pp. 948-959, 2005.

[27] C. Prigent, F. Aires, and W. B. Rossow, "Land surface microwave emissivities over the globe for a decade," *Bulletin of the American Meteorological Society,* vol. 87, no. 11, pp. 1573-1584, 2006.

[28] C. Prigent, E. Jaumouille, F. Chevallier *et al.*, "A parameterization of the microwave land surface emissivity between 19 and 100 GHz, anchored to satellite-derived estimates," *IEEE Transactions on Geoscience & Remote Sensing,* vol. 46, no. 2, pp. 344-352, 2008.

[29] P. de Rosnay, M. Drusch, A. Boone *et al.*, "AMMA land surface model intercomparison experiment coupled to the Community Microwave Emission Model: ALMIP-MEM," *Journal of Geophysical Research Atmospheres,* vol. 114, no. D05108, pp. 1-18, 2009.

[30] J. Moncet, P. Liang, A. Galantowicz *et al.*, "Land surface microwave emissivities derived from AMSR-E and MODIS measurements with advanced quality control," *Journal of Geophysical Research Atmospheres,* vol. 116, pp. D16104, 2011.

[31] S. Prakash, H. Norouzi, M. Azarderakhsh *et al.*, "Global land surface emissivity estimation from AMSR2 observations," *IEEE Geoscience and Remote Sensing Letters,* vol. 13, no. 9, pp. 1270-1274, 2016.

[32] Y. B. Qiu, J. C. Shi, M. T. Hallikainen *et al.*, "The AMSR-E instantaneous emissivity estimation and its correlation, frequency dependency analysis over different land covers," in IEEE International Geoscience & Remote Sensing Symposium, IGARSS 2008, Boston, Massachusetts, USA, 2009, pp. 749-752.

[33] H. Norouzi, M. Temimi, W. B. Rossow *et al.*, "The sensitivity of land emissivity estimates from AMSR-E at C and X bands to surface properties," *Hydrology and Earth System Sciences Discussions,* vol. 15, no. 11, pp. 3577-3589, 2011.

[34] J. H. Hu, Y. Y. Fu, P. Zhang *et al.*, "Satellite retrieval of microwave land surface emissivity under clear and cloudy skies in China using observations from AMSR-E and MODIS," *Remote Sensing,* vol. 13, pp. 3980, 2021.

[35] G. D. Wilke, and M. J. Mcfarland, "Correlations between Nimbus-7 scanning multichannel microwave radiometer data and an antecedent precipitation index," *Journal of Applied Meteorology and Climatology,* vol. 25, no. 2, pp. 227-238, 1986.

[36] Y. B. Qiu, J. C. Shi, M. Hallikainen *et al.*, "The AMSR-E instantaneous emissivity estimation and its correlation, frequency dependency analysis over different land covers," in In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Boston, MA, USA, 2008, pp. 749-752.

[37] F. J. Turk, Z. S. Haddad, and Y. You, "Principal components of multifrequency microwave land surface emissivities. Part I: Estimation under clear and precipitating conditions," *Journal of Hydrometeorology,* vol. 15, no. 1, pp. 3-19, 2014.

[38] C. S. Raju, T. Antony, N. Mathew *et al.*, "Mt-madras brightness temperature analysis for terrain characterization and land surface

microwave emissivity estimation," *Current Science,* vol. 104, pp. 1643-1649, 2013.

[39] Y. Wu, B. Qian, Y. S. Bao *et al.*, "Microwave land emissivity calculations over the Qinghai Tibetan plateau using FY-3BMWRI measurements," *Remote Sensing,* vol. 11, no. 19, pp. 2206, 2019.

[40] H. Yang, and F. Z. Weng, "Error sources in remote sensing of microwave land surface emissivity," *IEEE Transactions on Geoscience & Remote Sensing,* vol. 49, no. 9, pp. 3437-3442, 2011.

[41] Y. Tian, C. D. Peters-Lidard, K. W. Harrison *et al.*, "Quantifying uncertainties in land-surface microwave emissivity retrievals," *IEEE Transactions on Geoscience & Remote Sensing,* vol. 52, no. 2, pp. 829-840, 2014.

[42] S. Ringerud, C. Kummerow, C. Peters-Lidar *et al.*, "A comparison of microwave window channel retrieved and forward-modeled emissivities over the U.S. southern great plains," *IEEE Transactions on Geoscience & Remote Sensing,* vol. 52, no. 5, pp. 2395-2412, 2014.

[43] Y. Tian, C. D. Peters-Lidard, K. W. Harrison *et al.*, "An examination of methods for estimating land surface microwave emissivity," *Journal of Geophysical Research: Atmospheres,* vol. 120, no. 21, pp. 11114–111128, 2015.

[44] X. Cui, Z. Yao, Z. Zhao *et al.*, "Use of double channel differences for reducing the surface emissivity dependence of microwave atmospheric temperature and humidity retrievals," *Earth and Space Science,* vol. 7, no. 5, pp. 1-26, 2020.

[45] B. Lin, and P. Minnis, "Temporal variations of land surface microwave emissivities over the atmospheric radiation measurement program southern great plains site," *Journal of Applied Meteorology,* vol. 39, no. 7, pp. 1103-1116, 2000.

[46] F. Baordo, and A. J. Geer, "Assimilation of SSMIS humidity-sounding channels in all-sky conditions over land using a dynamic emissivity retrieval," *Quarterly Journal of the Royal Meteorological Society,* vol. 142, no. 700, pp. 2854-2866, 2016.

[47] B. Lin, B. Wielicki, P. Minnis *et al.*, "Estimation of water cloud properties from satellite microwave, infrared and visible measurements in oceanic environments: 1. Microwave brightness temperature simulations," *Journal of Geophysical Research,* vol. 103, no. D4, pp. 3873–3886, 1998.

[48] S. A. Clough, M. W. Shephard, E. J. Mlawer *et al.*, "Atmospheric radiative transfer modeling: A summary of the AER codes," *Journal of Quantitative Spectroscopy and Radiative Transfer,* vol. 91, no. 2, pp. 233-244, 2005.

[49] P. Ashcroft, and F. J. Wentz, "Algorithm Theoretical Basis Document (ATBD) AMSR Level 2A algorithm remote sensing systems: Santa Rosa, CA, USA," 2000.

[50] P. Ashcroft, and F. J. Wentz, "AMSR-E/Aqua L2A global swath spatially-resampled brightness temperatures; Version 3; NASA national snow and ice data center distributed active archive center: Boulder, CO, USA," 2013.

[51] Z. M. Wan, "New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product," *Remote Sensing of Environment,* vol. 140, pp. 36-45, 2014.

[52] S. B. Duan, Z. L. Li, H. Li *et al.*, "Validation of Collection 6 MODIS land surface temperature product using in situ measurements," *Remote Sensing of Environment,* vol. 225, pp. 16-29, 2019.

[53] X. M. Zhu, X. N. Song, P. Leng *et al.*, "Performances of six reanalysis profile products in the atmospheric correction of passive microwave data for estimating land surface temperature under cloudy-sky conditions," *International Journal of Digital Earth,* vol. 15, no. 1, pp. 296-322, 2022.

[54] H. Hersbach, B. Bell, P. Berrisford *et al.*, "The ERA5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society,* vol. 146, pp. 1999-2049, 2020.

[55] A. Chaieb, N. Rebair, and S. Bouazizb, "Vertical accuracy assessment of SRTM Ver 4.1 and ASTER GDEM Ver 2 using GPS measurements in central west of Tunisia," *Journal of Geographic Information System,* vol. 8, no. 1, pp. 57-64, 2016.

[56] F. C. Zhou, Z. L. Li, H. Wu *et al.*, "A remote sensing method for retrieving land surface emissivity and temperature in cloudy areas: a case study over South China," *International Journal of Remote Sensing,* vol. 40, no. 5-6, pp. 1724-1735, 2019.

[57] W. Snyder, and Z. M. Wan, "BRDF models to predict spectral reflectance and emissivity in the thermal infrared," *IEEE Transactions on Geoence & Remote Sensing,* vol. 36, no. 1, pp. 214-225, 1998.

[58] F. Z. Weng, and N. C. Grody, "Physical retrieval of land surface temperature using the special sensor microwave imager," *Journal of Geophysical Research: Atmospheres,* vol. 103, no. D8, pp. 8839-8848, 1998.

[59] S. J. English, "Airborne radiometric observations of cloud liquid-water emission at 89 and 157 GHz-application to retrieval of liquid-water path," *Quarterly Journal of the Royal Meteorological Society,* vol. 121, pp. 1501-1524, 1995.

[60] Y. C. Zhang, W. B. Rossow, and P. W. Stackhouse, "Comparison of different global information sources used in surface radiative flux calculation: radiative properties of the near surface atmosphere," *Journal of Geophysical Research-Atmospheres,* vol. 111, 2006.

[61] L. Breiman, "Random forests," *Machine Learning,* vol. 45, no. 1, pp. 5-32, 2001.

[62] C. Hutengs, and M. Vohland, "Downscaling land surface temperatures at regional scales with random forest regression," *Remote Sensing of Environment,* vol. 178, pp. 127-141, 2016.

[63] X. M. Zhu, X. N. Song, P. Leng *et al.*, "A framework for generating high spatiotemporal resolution land surface temperature in heterogeneous areas," *Remote Sensing,* vol. 13, no. 19, pp. 3885, 2021.

[64] H. L. Gao, R. Fu, R. E. Dickinson *et al.*, "A practical method for retrieving land surface temperature from AMSR-E over the Amazon forest," *IEEE Transactions on Geoence & Remote Sensing,* vol. 46, no. 1, pp. 193-199, 2008.

[65] E. Y. Zhao, C. X. Gao, X. G. Jiang *et al.*, "Land surface temperature retrieval from AMSR-E passive microwave data," *Optics Express,* vol. 25, no. 20, pp. 940-952, 2017.

[66] A. Mcgovern, R. Lagerquist, D. J. Gagne *et al.*, "Making the black box more transparent: Understanding the physical implications of machine learning," *Bulletin of the American Meteorological Society,* vol. 100, no. 1, pp. 2175-2199, 2019.

[67] R. R. Ferraro, C. D. Peters-Lidard, C. Hernandez *et al.*, "An evaluation of microwave land surface emissivities over the continental United States to benefit GPM-Era precipitation algorithms," *IEEE Transactions on Geoscience & Remote Sensing,* vol. 51, no. 1, pp. 378-398, 2013.

[68] H. Norouzi, W. Rossow, M. Temimi *et al.*, "Using microwave brightness temperature diurnal cycle to improve emissivity retrievals over land," *Remote Sensing of Environment,* vol. 123, pp. 470-482, 2012.

[69] J. W. Salisbury, and D. M. D'Aria, "Emissivity of terrestrial materials in the 8-14 μm atmospheric window," *remote Sensing of Environment,* vol. 42, pp. 83-106, 1992.

[70] W. G. Rees, "Infrared emissivities of Arctic land cover types," *International Journal of Remote Sensing,* vol. 14, pp. 1013-1017, 1993.

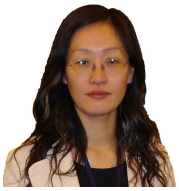
[71] T. J. Jackson, and P. E. O'Neill, "Salinity effects on the microwave emission of soil," *IEEE Transactions on Geoence & Remote Sensing,* vol. 25, pp. 214-220, 1987.

[72] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive, Volume III, from Theory to Applications*, North Bergen, NJ: Book-Mart Press, Inc, 1986.

[73] H. Norouzi, M. Temimi, C. Prigent *et al.*, "Assessment of the consistency among global microwave land surface emissivity products," *Atmospheric Measurement Techniques Discussions,* vol. 7, no. 9, pp. 9993-10013, 2015.

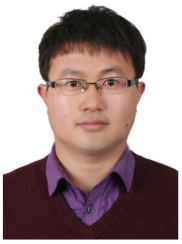
Xin-Ming Zhu received the M.S. degree from the Northwest University, Xi'an, China, in 2018. He is currently working toward the Ph.D. degree at the University of Chinese Academy of Sciences, Beijing.

His current research interests include the retrieval and spatio-temporal analysis of land surface temperature and land surface emissivity from passive microwave remote sensing.



Xiao-Ning Song received the Ph.D. degree in cartography and geographical information system from Graduated the University of Chinese Academy of Sciences, Beijing, China, in 2004.

She is currently a Professor with the University of Chinese Academy of Sciences. Her current research interests include quantitative remote sensing and the application of remote sensing in ecological environment.



Pei Leng received the Ph.D. degree in cartography and geographical information system from Graduated the University of Chinese Academy of Sciences, Beijing, China, in 2015.

He is currently an Associate Professor with the Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing. His current research interests include the retrieval and validation of soil moisture content.

Zhao-Liang Li received the Ph.D. degree in remote sensing from Louis Pasteur University (currently called University of Strasbourg), Strasbourg, France, in 1990.

Since 1992, he has been a Research Scientist with the National Center for Scientific Research (CNRS), Illkirch, France. In 2012, he joined the Institute of Agricultural Resources and Regional Planning. He has participated in many national and international projects, such as the NASA-funded MODIS, the EC-funded program EAGLE, and the ESA-funded program SPECTRA. He has authored more than 200 papers in international refereed journals. His research interests include thermal infrared radiometry, parameterization of land surface processes at large scale, and the assimilation of satellite data to land surface models.

Xiao-Tao Li received the M.S. degree in Mineral Resource Prospecting and Exploration from Graduated the Shandong University of Science and Technology, Qingdao, China, in 2004.

He is currently a Professor with the Chinese Institute of Water Resource and Hydropower Research. His current research interests include the drought monitoring and flood monitoring using remote sensing image.

Liang Gao received the M.S. degree from the Aerospace Information Research Institute, Chinese Academy of Science, Beijing, in 2020. He is currently working toward the Ph.D. degree at the University of Chinese Academy of Sciences, Beijing.

His current research interests include the retrieval and validation of soil moisture content using active microwave remote sensing.

Da Guo received the B.E. degree from the Shandong University of Science and Technology, Qingdao, China, in 2018. He is currently working toward the Ph.D. degree at the University of Chinese Academy of Sciences, Beijing.

His current research interests include the retrieval and validation of leaf area index using laser radar remote sensing.